Mathematical Programming
https://doi.org/10.1007/s10107-021-01657-8

FULL LENGTH PAPER

Series B

Sparse representation of vectors in lattices and semigroups

Iskander Aliev[1] · Gennadiy Averkov[2] · Jesús A. De Loera[3] · Timm Oertel[1]

Received: 30 June 2020 / Accepted: 21 April 2021

Abstract

We study the sparsity of the solutions to systems of linear Diophantine equations with and without non-negativity constraints. The sparsity of a solution vector is the number of its nonzero entries, which is referred to as the ℓ_0-norm of the vector. Our main results are new improved bounds on the minimal ℓ_0-norm of solutions to systems $A\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbb{Z}^{m\times n}$, $\boldsymbol{b} \in \mathbb{Z}^m$ and $\boldsymbol{x}$ is either a general integer vector (lattice case) or a non-negative integer vector (semigroup case). In certain cases, we give polynomial time algorithms for computing solutions with ℓ_0-norm satisfying the obtained bounds. We show that our bounds are tight. Our bounds can be seen as functions naturally generalizing the rank of a matrix over $\mathbb{R}$, to other subdomains such as $\mathbb{Z}$. We show that these new rank-like functions are all NP-hard to compute in general, but polynomial-time computable for fixed number of variables.

Mathematics Subject Classification 20M10 · 52C07 · 90C10 · 94A12

Earlier conference proceeding's version: I. Aliev, G. Averkov, J. A. De Loera and T. Oertel, *Optimizing Sparsity over Lattices and Semigroups*, IPCO 2020, LNCS 12125, (2020) pp. 40–51.

✉ Timm Oertel
oertelt@cardiff.ac.uk

Iskander Aliev
alievi@cardiff.ac.uk

Gennadiy Averkov
averkov@b-tu.de

Jesús A. De Loera
deloera@math.ucdavis.edu

1 Cardiff University, Cardiff, UK

2 BTU Cottbus - Senftenberg, Cottbus, Germany

3 University of California, Davis, Davis, USA

Published online: 04 May 2021

1 Introduction

Given a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^m$, we study the sparsity of solutions to the system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$, in variables $x_1, \ldots, x_n$ restricted to a structured domain $D \subseteq \mathbb{R}$. The sparsity of these solutions is quantified via the ℓ_0-*norm*, which is the size $\|\boldsymbol{x}\|_0 := |\operatorname{supp}(\boldsymbol{x})|$ of the *support* $\operatorname{supp}(\boldsymbol{x}) := \{i \,:\, x_i \neq 0\}$ of the vector $\boldsymbol{x}$. The sparsest solutions are optimal solutions of the optimization problem

$$\min\left\{\|\boldsymbol{x}\|_0 \,:\, A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in D^n\right\}. \tag{1}$$

When $D = \mathbb{R}$ and $D = \mathbb{R}_{\geq 0}$, the tight upper bounds on (1) in terms of A are given by the rank of A, which follows from basic linear algebra and the well-known Carathéodory's theorem from convexity, respectively. Nevertheless, even when $D = \mathbb{R}$, computation of (1) for given A and $\boldsymbol{b}$ is NP-hard [26].

The ℓ_0-norm minimization problem (1) is central in the theory of the compressed sensing, where for the classical choice $D = \mathbb{R}$ an appropriate linear programming relaxation of (1) provides a guaranteed approximation [9,11,12]. In the present paper, we deal with the two discrete domains, $D = \mathbb{Z}$ and $D = \mathbb{Z}_{\geq 0}$, which are naturally related to the theory of systems of linear Diophantine equations and integer linear programming, respectively.

Sparsity of solutions to linear Diophantine equations is relevant for the theory of compressed sensing for integer-valued signals [17,18,24], motivated by many applications in which the signal is known to have integer entries, for instance, in wireless communication [31] and in the theory of error-correcting codes [10]. Support minimization was also investigated in connection to integer optimization [2,16,29,30]. Also, numerous applications to combinatorial optimization problems have been explored. For example, the minimum edge-coloring problem can be seen as finding the sparsest representation in the semigroup generated by the matchings of the graph [13,25]. Further examples of combinatorial applications can be found in [4] and [16].

Since we know that for $D = \mathbb{R}$, the sparsity of solution is captured by the notion of the rank of A, we introduce a similar notion with respect to an arbitrary underlying domain D. We define the D-*rank* of A as

$$\max_{\boldsymbol{y}\in D^n} \min\{\|\boldsymbol{x}\|_0 : A\boldsymbol{x} = A\boldsymbol{y}, \boldsymbol{x} \in D^n\}.$$

In this respect, note that the computation of the D-rank is a *bi-level optimization* problem. There is yet another natural generalization of the notion of rank in our setting. For that let $[n] := \{1, \ldots, n\}$, let $\binom{[n]}{k}$ be the set of all k-element subsets of $[n]$, and for $\gamma \in \binom{[n]}{k}$ let A_γ denote the $m \times k$ submatrix of A with columns indexed by γ. Then the D-*complexity* of A is defined as the minimal $k \in \mathbb{Z}_{\geq 0}$ such that there exists a $\tau \in \binom{[n]}{k}$ for which the following equality holds $\{A\boldsymbol{x} \,:\, \boldsymbol{x} \in D^n\} = \left\{A_\tau \boldsymbol{y} \,:\, \boldsymbol{y} \in D^k\right\}$. It is clear that the D-rank is bounded from above by the D-complexity.

As examples, note that when the domain $D = \mathbb{R}$, both D-rank and D-complexity coincide with the rank of the matrix A from linear algebra. For $D = \mathbb{R}_{\geq 0}$, the D-rank

is again the regular rank of A, but the D-complexity is in general larger. If the columns of A positively span a pointed cone, then the D-complexity is the number of extreme rays of this cone.

In this paper we specialize the above two functions to the two domains $D = \mathbb{Z}$ and $D = \mathbb{Z}_{\geq 0}$, which yields a natural geometric interpretation in terms of lattices and semigroups. First, the matrix A determines the *lattice* $\mathcal{L}(A) := \{A\boldsymbol{x} : \boldsymbol{x} \in \mathbb{Z}^n\}$ generated by the columns of A. Secondly, the matrix A determines the *semigroup* $\mathcal{S}\}(A) := \{A\boldsymbol{x} : \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}\}$. Note that this set consists of all right-hand-side vectors $\boldsymbol{b}$, for which the system $A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$ of integer-programming constraints on $\boldsymbol{x}$ is feasible. We obtain the following four functions:

ILR(A) (*integer linear rank*): The $\mathbb{Z}$-rank of A, i.e., the minimal k s.t. $\mathcal{L}(A) = \bigcup_{\tau \in \binom{[n]}{k}} \mathcal{L}(A_\tau)$.

ICR(A) (*integer Carathéodory rank*): The $\mathbb{Z}_{\geq 0}$-rank of A, i.e., the minimal k s.t. $\mathrm{Sg}(A) = \bigcup_{\tau \in \binom{[n]}{k}} \mathrm{Sg}(A_\tau)$

ILC(A) (*integer linear complexity*): The $\mathbb{Z}$-complexity of A, i.e., the minimal k s.t. $\mathcal{L}(A) = \mathcal{L}(A_\tau)$ holds for some $\tau \in \binom{[n]}{k}$.

ICC(A) (*integer Carathéodory complexity*): The $\mathbb{Z}_{\geq 0}$-complexity of A, i.e., the minimal k s.t. $\mathrm{Sg}(A) = \mathrm{Sg}(A_\tau)$ for some $\tau \in \binom{[n]}{k}$.

In our results we deal with an integer matrix $A \in \mathbb{Z}^{m \times n}$, and, without loss of generality, A is assumed to have a full row rank. In this case, we have that the determinant of the lattice $\mathcal{L}(A)$ is equal to

$$\gcd(A) := \gcd\left\{\det(A_\gamma) : \gamma \in \binom{[n]}{m}\right\}.$$

See for example [30, Section 1.3]. For a general introduction to lattices see [21].

1.1 Bounds for ILR(A) and ILC(A)

For stating our results, we need several number-theoretic functions. Given $z \in \mathbb{Z}_{>0}$, consider the prime factorization $z = p_1^{s_1} \cdots p_k^{s_k}$ with pairwise distinct prime factors $p_1, \ldots, p_k$ and their multiplicities $s_1, \ldots, s_k \in \mathbb{Z}_{>0}$. Then the number of prime factors $\sum_{i=1}^k s_i$ counting the multiplicities is denoted by $\Omega(z)$. Furthermore, we introduce

$$\Omega_m(z) := \sum_{i=1}^{k} \min\{s_i, m\}.$$

That is, by introducing m we set a threshold to account for multiplicities. In the case $m = 1$ we thus have

$$\omega(z) := \Omega_1(z) = k,$$

which is the number of prime factors in z, not taking the multiplicities into account. The functions Ω and ω are called *prime Ω-function* and *prime ω-function*, respectively, in number theory [23]. We call Ω_m the *truncated prime Ω-function*.

Theorem 1 *Let $A \in \mathbb{Z}^{m\times n}$ be a matrix of rank m. Then*

$$\mathit{ILR}(A) \leq \mathit{ILC}(A) \leq m + \min_{\substack{\tau\in\binom{[n]}{m}\\ \det(A_\tau)\neq 0}} \Omega_m\left(\frac{|\det(A_\tau)|}{\gcd(A)}\right). \tag{2}$$

One can easily see that the estimates $\omega(z) \leq \Omega_m(z) \leq \Omega(z) \leq \log_2(z)$ hold for every $z \in \mathbb{Z}_{>0}$. The estimate using $\log_2(z)$ gives a first impression on the size of the bound (2). It turns out, however, that $\Omega_m(z)$ is much smaller on the average. Results in number theory [23, §22.10] show that the average values $\frac{1}{z}(\omega(1)+\cdots+\omega(z))$ and $\frac{1}{z}(\Omega(1)+\cdots+\Omega(z))$ are of order $\log(\log(z))$, as $z \to \infty$.

In Proposition 1 from Sect. 5, we show that (2) is an optimal bound on both $\mathsf{ILR}(A)$ and $\mathsf{ILC}(A)$, in the sense that neither m can be replaced by any smaller constant nor the function Ω_m occurring on the right-hand side can be replaced by any smaller function. Furthermore, as a byproduct of our constructive proof of Theorem 1, we obtain the following.

Corollary 1 *Let $A \in \mathbb{Z}^{m\times n}$ be a matrix of full row rank and let A_τ be a non-singular sub-matrix of A, where $\tau \in \binom{[n]}{m}$. If the system $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n$ with a right-hand side $\boldsymbol{b} \in \mathbb{Z}^m$ is feasible, then this system has a solution $\boldsymbol{x}$ that satisfies*

$$\|\boldsymbol{x}\|_0 \leq m + \Omega_m\left(\frac{|\det(A_\tau)|}{\gcd(A)}\right).$$

For the input A, τ and $\boldsymbol{b}$ in binary encoding, such a solution $\boldsymbol{x}$ can be computed in polynomial time.

1.2 Bounds for ICR(A) and ICC(A)

Theorem 1.1(i) in [3] (see also [2, Theorem 1]) immediately implies the bound

$$\mathsf{ICR}(A) \leq m + \left\lfloor \log_2\left(\frac{\sqrt{\det(AA^\top)}}{\gcd(A)}\right)\right\rfloor, \tag{3}$$

which is an improvement of [16, Theorem 1(ii)]. Note that

$$\det(AA^\top) = \sum_{\tau\in\binom{[n]}{m}} \det(A_\tau)^2 \tag{4}$$

by the Cauchy–Binet formula.

We show that, under natural assumptions on A, we can significantly improve this bound. First, we consider matrices A whose columns positively span $\mathbb{R}^m$. Theorem 1 can be used in this case to obtain the following result.

Theorem 2 *Let $A \in \mathbb{Z}^{m\times n}$ be a matrix whose columns positively span $\mathbb{R}^m$. Then*

$$ICR(A) \le ICC(A) \le 2m + \min_{\substack{\tau\in\binom{[n]}{m}\\ \det(A_\tau)\neq 0}} \Omega_m\left(\frac{|\det(A_\tau)|}{\gcd(A)}\right). \tag{5}$$

Both the general bound (3) and our bound (5) have the first term linearly depending on m and the second term depending on the $m \times m$ minors of A scaled by $\gcd(A)$. Thus, taking into account $|\det(A_\tau)| \le \sqrt{\det(AA^\top)}$ and $\Omega_m(z) \le \log_2(z)$, we see that the second term in (5) is not larger than the second term in the bound (3): in fact, the second term in (5) is much smaller "on the average". As for (2), we show in Proposition 2 from Sect. 5 that, under the given assumptions on A, the bound (5) is optimal.

In the knapsack case $m = 1$, the bound (5) strengthens Theorem 1.2 in [3] and, as it was already indicated in the IPCO version of this paper [1], it confirms a conjecture posed in [3, page 247]. Moreover, as a byproduct of the proof of Theorem 2 we obtain the following algorithmic result.

Corollary 2 *Let $A \in \mathbb{Z}^{m\times n}$ be a matrix whose columns positively span $\mathbb{R}^m$ and let A_τ be an $m \times m$ non-singular sub-matrix of A, where $\tau \in \binom{[n]}{m}$. If the feasibility problem $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n_{\ge 0}$ of integer programming with a right-hand side $\boldsymbol{b} \in \mathbb{Z}^m$ has a solution, then it has a solution satisfying*

$$\|\boldsymbol{x}\|_0 \le 2m + \Omega_m\left(\frac{|\det(A_\tau)|}{\gcd(A)}\right).$$

For the input A, τ and $\boldsymbol{b}$ in binary encoding, such a solution $\boldsymbol{x}$ can be computed in polynomial time.

Our next contribution gives an improvement on (3) for the case when the columns of A generate a pointed cone. Given $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$, we denote by $\mathrm{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ the convex conic hull of the set $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$. Assume that the matrix $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{Z}^{m\times n}$ with columns $\boldsymbol{a}_i$ satisfies the following conditions:

$$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{Z}^m \setminus \{\boldsymbol{0}\}, \tag{6}$$

$$\mathrm{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \text{ is an } m\text{-dimensional pointed cone}, \tag{7}$$

$$\mathrm{cone}(\boldsymbol{a}_1) \text{is an extreme ray of } \mathrm{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n). \tag{8}$$

Theorem 3 *Let $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{Z}^{m\times n}$ satisfy (6)–(8). Then*

$$ICR(A) \le m + \left\lfloor \log_2\left(\frac{q(A)}{\gcd(A)}\right)\right\rfloor, \tag{9}$$

where

$$q(A) := \sqrt{\sum_{I \in \binom{[n]}{m} : 1 \in I} \det(A_I)^2}. \tag{10}$$

In view of (4), the bound (9) improves on (3) by reducing the sum over all $I \in \binom{[n]}{m}$ to the sum over those I that satisfy $1 \in I$. The proof of Theorem 3 will be derived as an extension of the proof of our next result in the setting of the knapsack scenario $m = 1$. In this setting, $A = \boldsymbol{a}$ is a row vector and the assumption (7) is equivalent to $\boldsymbol{a} \in \mathbb{Z}_{>0}^{1\times n} \cup \mathbb{Z}_{<0}^{1\times n}$. Without loss of generality, one can assume $\boldsymbol{a} \in \mathbb{Z}_{>0}^{1\times n}$.

Theorem 4 *Let* $\boldsymbol{a} = (a_1, \ldots, a_n) \in \mathbb{Z}_{>0}^{1\times n}$. *Then*

$$ICR(\boldsymbol{a}) \leq 1 + \left\lfloor \log_2 \left(\frac{\min\{a_1, \ldots, a_n\}}{\gcd(\boldsymbol{a})} \right) \right\rfloor. \tag{11}$$

Similar to (5), we show in Proposition 3 from Sect. 5 that the bounds (9) and (11) are optimal.

1.3 Computational complexity

It is well known that the feasibility problem in integer linear programming is NP-complete (see [32, § 18.2]), which means that testing whether the sparsity optimization problem (1) is feasible is hard in the case $D = \mathbb{Z}_{\geq 0}^n$. But even in the cases when testing feasibility is tractable, solving (1) is usually hard due to the hardness of the ℓ_0-norm as an objective. For example, computation of (1) is NP-hard for $D = \mathbb{R}^n$ (see [26]). In Sect. 1.3 we study the complexity of computing our four rank-like functions $\mathsf{ILR}(A)$, $\mathsf{ILC}(A)$, $\mathsf{ICR}(A)$ and $\mathsf{ICC}(A)$. We would like to emphasize that the complexity analysis of these functions is more intricate than the respective analysis of (1).

Theorem 5 *Consider the four problems of verifying* $\mathsf{ILC}(A) \leq k$, $\mathsf{ILR}(A) \leq k$, $\mathsf{ICR}(A) \leq k$, *and* $\mathsf{ICC}(A) \leq k$, *for given* $A \in \mathbb{Z}^{m\times n}$ *and* $k \in \mathbb{Z}_{>0}$. *These problems have the following complexity:*

(i) $\mathsf{ILR}(A) \leq k$, $\mathsf{ILC}(A) \leq k$, *and* $\mathsf{ICC}(A) \leq k$ *are NP-complete when* $m = 1$ *and* n *is a part of the input,*
(ii) $\mathsf{ICR}(A) \leq k$ *is NP-hard when* $m = 1$ *and* n *is a part of the input,*
(iii) $\mathsf{ILC}(A) \leq k$ *and* $\mathsf{ICC}(A) \leq k$ *are strongly NP-complete when* m *and* n *are a part of the input,*
(iv) $\mathsf{ILR}(A) \leq k$ *and* $\mathsf{ICR}(A) \leq k$ *are strongly NP-hard when* m *and* n *are a part of the input.*

In the case $m = 1$, one might be tempted to compare (1) for $D = \mathbb{Z}^n$ and $D = \mathbb{Z}_{\geq 0}^n$ with the integer knapsack optimization problem. For the latter, the linearity of the objective function allows one to use dynamic programming to solve the problem in pseudo-polynomial time. It is however not clear if it is possible to adapt this approach for the case of the ℓ_0 objective. This motivates the following problem:

Problem 1 If $m = 1$, can (1) and the functions $\mathsf{ILR}(A)$, $\mathsf{ILC}(A)$, $\mathsf{ICR}(A)$ and $\mathsf{ICC}(A)$ be computed in pseudo-polynomial time, i.e., when the input numbers are given in unary encoding rather than binary?

Finally, we want to address the case when the number of variables n is fixed. It is easy to see that the optimization problem (1) can be solved in polynomial time for both $D = \mathbb{Z}^n$ and $D = \mathbb{Z}^n_{\geq 0}$. For a fixed n, all 2^n possible choices of the support for the vector $\boldsymbol{x}$ can be enumerated. For each such choice, the existence of the vector $\boldsymbol{x}$ with $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in D^n$ and the prescribed support can be checked in polynomial time: for $D = \mathbb{Z}^n$ one needs so solve a Diophantine system, while for $D = \mathbb{Z}^n_{\geq 0}$ one uses polynomial-time solvability of integer linear problems in fixed dimension [32, § 18.4]. Rather similarly, one can also establish polynomial-time solvability of $\mathsf{ILC}(A)$ and $\mathsf{ICC}(A)$ for a fixed n. In contrast to this, since the D-ranks are related to bi-level programming, the study of the computational complexity of $\mathsf{ILR}(A)$ and $\mathsf{ICR}(A)$ in the case of fixed n requires a more involved algorithmic theory, which has been developed only recently. Using recent results in the algorithmic theory of Presburger arithmetic, we obtain the following.

Theorem 6 *When n is fixed, for given $m \in \mathbb{Z}_{>0}$ and $A \in \mathbb{Z}^{m\times n}$, all four values $\mathsf{ILR}(A)$, $\mathsf{ICR}(A)$, $\mathsf{ILC}(A)$ and $\mathsf{ICC}(A)$ can be computed in polynomial time.*

2 Proofs of Theorem 1 and Corollary 1

The proof of Theorem 1 relies on the theory of finite Abelian groups (see [15] for a general reference). We write Abelian groups additively. An Abelian group G is a *direct sum* of its finitely many subgroups $G_1, \ldots, G_m$, which is written as $G = \bigoplus_{i=1}^m G_i$, if every element $x \in G$ has a unique representation as $x = x_1 + \cdots + x_m$ with $x_i \in G_i$ for each $i \in [m]$. A *primary cyclic group* is a non-zero finite cyclic group whose order is a power of a prime number. We use G/H to denote the quotient of G modulo its subgroup H.

The fundamental theorem of finite Abelian groups states that every finite Abelian group G has a *primary decomposition*, which is essentially unique. This means, G is decomposable into a direct sum of its primary cyclic subgroups and that this decomposition is unique up to automorphisms of G (see Theorems 3 and 5 in Chapter 5.2 of [15], with further details in 12.1). We denote by $\kappa(G)$ the number of direct summands in the primary decomposition of G.

For a subset S of a finite Abelian group G, we denote by $\langle S\rangle$ the subgroup of G generated by S. We call a subset S of G *non-redundant* if the subgroups $\langle T\rangle$ generated by proper subsets T of S are properly contained in $\langle S\rangle$. The following result gives an upper bound on the maximum cardinality of S.

Theorem 7 *Let G be a finite Abelian group. Then the maximum cardinality of a non-redundant subset S of G is equal to $\kappa(G)$.*

Even though this result is available in the literature (see, for example, [20, Lemma A.6]), it does not seem to be well known, and we have not found any source

containing a complete self-contained proof of this result. Thus, we provide a proof of Theorem 7 in the Appendix, relying only on the basic facts from group theory. We will also need the following lemmas.

Lemma 1 *Let G be a finite Abelian group representable as a direct sum $G = \bigoplus_{j=1}^{m} G_j$, where the groups $G_1, \ldots, G_m$ are cyclic. Then $\kappa(G) \leq \Omega_m(|G|)$.*

Proof Let us consider the prime factorization $|G| = p_1^{n_1} \cdots p_s^{n_s}$. Then $|G_j| = p_1^{n_{i,j}} \cdots p_s^{n_{i,j}}$ with $0 \leq n_{i,j} \leq n_i$ and, by the Chinese Remainder Theorem, the cyclic group G_j can be represented as $G_j = \bigoplus_{i=1}^{s} G_{i,j}$, where $G_{i,j}$ is a cyclic group of order $p_i^{n_{i,j}}$. Consequently, $G = \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{m} G_{i,j}$. This is a decomposition of G into a direct sum of primary cyclic groups and, possibly, some trivial summands $G_{i,j}$ equal to $\{0\}$. We can count the non-trivial direct summands whose order is a power of p_i, for a given $i \in [s]$. There is at most one summand like this for each of the groups G_j. So, there are at most m non-trivial summands in the decomposition whose order is a power of p_i. On the other hand, the direct sum of all non-trivial summands whose order is a power of p_i is a group of order $p_i^{n_{i,1}+\cdots+n_{i,s}} = p_i^{n_i}$ so that the total number of such summands is not larger than n_i, as every summand contributes the factor at least p_i to the power $p_i^{n_i}$. Thus, the total number of non-zero summands in the decomposition of G is at most $\sum_{i=1}^{s} \min\{m, n_i\} = \Omega_m(|G|)$. □

Lemma 2 *Let Λ be a m–dimensional sublattice of $\mathbb{Z}^m$. Then $G = \mathbb{Z}^m/\Lambda$ is a finite Abelian group of order* $\det(\Lambda)$ *that can be represented as a direct sum of at most m cyclic groups.*

Proof The proof relies on the relationship of finite Abelian groups and lattices, see [32, §4.4]. Fix a matrix $M \in \mathbb{Z}^{m\times m}$ whose columns form a basis of Λ. Then $|\det(M)| = \det(\Lambda)$. There exist unimodular matrices $U \in \mathbb{Z}^{m\times m}$ and $V \in \mathbb{Z}^{m\times m}$ such that $D := UMV$ is diagonal matrix with positive integer diagonal entries. For example, one can choose D to be the Smith Normal Form of M [32, §4.4]. Let $d_1, \ldots, d_m \in \mathbb{Z}_{>0}$ be the diagonal entries of D. Since U and V are unimodular, $d_1 \cdots d_m = \det(D) = \det(\Lambda)$.

We introduce the quotient group $G' := \mathbb{Z}^m/\Lambda' = (\mathbb{Z}/d_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/d_m\mathbb{Z})$ with respect to the lattice $\Lambda' := \mathcal{L}(D) = (d_1\mathbb{Z}) \times \cdots \times (d_m\mathbb{Z})$. The order of G' is $d_1 \cdots d_m = \det(D) = \det(\Lambda)$ and G' is a direct sum of at most m cyclic groups, as every $d_i > 1$ determines a non-trivial direct summand.

To conclude the proof, it suffices to show that G' is isomorphic to G. To see this, note that $\Lambda' = \mathcal{L}(D) = \mathcal{L}(UMV) = \mathcal{L}(UM) = \{Uz : z \in \Lambda\}$. Thus, the map $z \mapsto Uz$ is an automorphism of $\mathbb{Z}^m$ and an isomorphism from Λ to Λ'. Thus, $z \mapsto Uz$ induces an isomorphism from the group $G = \mathbb{Z}^m/\Lambda$ to the group $G' = \mathbb{Z}^m/\Lambda'$. □

The following lemma allows us to reduce considerations to the case $\gcd(A) = 1$, without affecting the sparsity.

Lemma 3 *Let $A \in \mathbb{Z}^{m\times n}$ have a full row rank and let $M \in \mathbb{Z}^{m\times m}$ be a matrix whose columns form a basis for $\mathcal{L}(A)$. Then the following holds:*

(a) $M^{-1}A$ *is an integer matrix of full row rank satisfying* $\gcd(M^{-1}A) = 1$.
(b) *The map* $\boldsymbol{b} \mapsto M^{-1}\boldsymbol{b}$, *as a map from* $\mathcal{L}(A)$ *to* $\mathcal{L}(M^{-1}A) = \mathbb{Z}^m$, *is an isomorphism of lattices, and, as a map from* $\mathrm{Sg}(A)$ *to* $\mathrm{Sg}(M^{-1}A)$, *is an isomorphism of the semigroups.*
(c) $\mathit{ILC}(A) = \mathit{ILC}(M^{-1}A)$, $\mathit{ILR}(A) = \mathit{ILR}(M^{-1}A)$, $\mathit{ICC}(A) = \mathit{ICC}(M^{-1}A)$, *and* $\mathit{ICR}(A) = \mathit{ICR}(M^{-1}A)$.
(d) *For given* m, n *and* A, *a matrix* M *as above can be computed in polynomial time.*

Proof (a) follows from $\gcd(M^{-1}A) = \gcd(A)/|\det(M)| = \gcd(A)/\det(\mathcal{L}(A)) = 1$, (b) is straightforward, and (c) follows from (b). To show (d), we observe that M can be obtained from the Hermite Normal Form H of A (with respect to the column transformations) by discarding the zero columns of H. □

Lemma 4 *Let* $A \in \mathbb{Z}^{m\times n}$ *have a full row rank and let* A_τ *be a non-singular* $m \times m$ *minor of* A, *where* $\tau \in \binom{[n]}{m}$. *Then there exists* $\gamma \subseteq [n]$ *with* $\tau \subseteq \gamma$ *such that* $\mathcal{L}(A_\gamma) = \mathcal{L}(A)$ *and* $|\gamma| \le m + \Omega_m\left(\frac{|\det(A_\tau)|}{\gcd(A)}\right)$. *For the input* A *and* τ *in binary encoding, such* γ *can be computed in polynomial time.*

Proof By Lemma 3, it suffices to consider the case $\gcd(A) = 1$. Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ be the columns of A. Without loss of generality, let $\tau = [m]$.

The matrix A_τ gives rise to the lattice $\Lambda := \mathcal{L}(A_\tau)$ of rank m, while Λ determines the finite Abelian group $\mathbb{Z}^m/\Lambda$. Consider the canonical homomorphism $\phi : \mathbb{Z}^m \to \mathbb{Z}^m/\Lambda$, sending an element of $\mathbb{Z}^m$ to its coset modulo Λ. Since $\gcd(A) = 1$, we have $\mathcal{L}(A) = \mathbb{Z}^m$, which implies $\langle T\rangle = \mathbb{Z}^m/\Lambda$ for $T := \{\phi(\boldsymbol{a}_{m+1}), \ldots, \phi(\boldsymbol{a}_n)\}$. For every non-redundant subset S of $\mathbb{Z}^m/\Lambda$ or $\langle T\rangle$, we have

$$\begin{aligned} |S| &\le \kappa(\mathbb{Z}^m/\Lambda) && \text{(by Theorem 7)} \\ &\le \Omega_m(|\det(A_\tau)|) && \text{(by Lemmas 1 and 2).} \end{aligned}$$

We fix a set $I \subseteq \{m+1, \ldots, n\}$ that satisfies $|I| = |S|$ and $S = \{\phi(\boldsymbol{a}_i) : i \in I\}$ and introduce

$$\gamma = I \cup \tau\,. \tag{12}$$

In this notation, equality $\langle T\rangle = \langle S\rangle = \mathbb{Z}^m/\Lambda$ can be reformulated as $\mathbb{Z}^m = \mathcal{L}(A_I) + \Lambda = \mathcal{L}(A_I) + \mathcal{L}(A_\tau) = \mathcal{L}(A_\gamma)$.

Let us now show that γ can be determined in polynomial time. It is enough to determine the set I, which defines the non-redundant subset $S = \{\phi(\boldsymbol{a}_i) : i \in I\}$ of $\mathbb{Z}^m/\Lambda$. Start with $I = \{m+1, \ldots, n\}$ and iteratively check if some of the elements $\phi(\boldsymbol{a}_i) \in \mathbb{Z}^m/\Lambda$, where $i \in I$, is in the group generated by the remaining elements. Suppose $j \in I$ and we want to check if $\phi(\boldsymbol{a}_j)$ is in the group generated by all $\phi(\boldsymbol{a}_i)$ with $i \in I\backslash\{j\}$. Since $\Lambda = \mathcal{L}(A_\tau)$, this is equivalent to checking $\boldsymbol{a}_j \in \mathcal{L}(A_{I\backslash\{j\}\cup\tau})$ and is thus reduced to solving a system of linear Diophantine equations with the left-hand side matrix $A_{I\backslash\{j\}\cup\tau}$ and the right-hand side vector $\boldsymbol{a}_j$ (such systems can be solved in polynomial time by [32, Corollary 5.3b]). Thus, carrying out the above procedure for every $j \in I$ and removing j from I whenever $\boldsymbol{a}_j \in \mathcal{L}(A_{I\backslash\{j\}\cup\tau})$, we eventually arrive at a set I that determines a non-redundant subset S of $\mathbb{Z}^m/\Lambda$. This is done by solving

at most $n - m$ linear Diophantine systems in total, where the matrix of each system is a sub-matrix of A and the right-hand side vector of the system is a column of A. □

Proof of Theorem 1 Consider an arbitrary $\tau \in \binom{[n]}{m}$, for which the matrix A_τ is non-singular, and the respective γ as in Lemma 4. One has $\mathsf{ILC}(A) \leq |\gamma| \leq m + \Omega_m \left(\frac{|\det(A_\tau)|}{\gcd(A)} \right)$. This yields (2). □

Proof of Corollary 1 We use γ from Lemma 4. Since $\boldsymbol{b} \in \mathcal{L}(A) = \mathcal{L}(A_\gamma)$, there exists a solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n$ with $\mathrm{supp}(\boldsymbol{x}) \subseteq \gamma$. This solution can be computed by solving the Diophantine system with the left-hand side matrix A_γ and the right-hand side vector $\boldsymbol{b}$. □

3 Proofs of Theorem 2 and Corollary 2

Lemma 5 *Let $A \in \mathbb{Z}^{m \times n}$ be a matrix whose columns positiely span $\mathbb{R}^m$ and A_τ be a non-singular $m \times m$ sub-matrix of A, where $\tau \in \binom{[n]}{m}$. Then there exists $I \subseteq [n]$ satisfying $\tau \subseteq I$, $|I| \leq 2m + \Omega_m \left(\frac{|\det(A_\tau)|}{\gcd(A)} \right)$ and $\mathcal{L}(A) = \mathcal{L}(A_I)$ such that the columns of A_I positively span $\mathbb{R}^m$. For the input A and τ in binary encoding, such I can be computed in polynomial time.*

Proof Consider γ as in Lemma 4. Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ be the columns of A. Since the matrix A_τ is non-singular, the m vectors $\boldsymbol{a}_i$, where $i \in \tau$, together with the vector $\boldsymbol{v} = -\sum_{i \in \tau} \boldsymbol{a}_i$ positively span $\mathbb{R}^n$. Since all columns of A positively span $\mathbb{R}^n$, the conic version of Carathéodory's theorem implies the existence of a set $\beta \subseteq [n]$ with $|\beta| \leq m$, such that $\boldsymbol{v}$ is in the conic hull of $\{\boldsymbol{a}_i : i \in \beta\}$. Consequently, the set $\{\boldsymbol{a}_i : i \in \beta \cup \tau\}$ and, by this, also the larger set $\{\boldsymbol{a}_i : i \in \beta \cup \gamma\}$ positively span $\mathbb{R}^m$. Thus, in view of Lemma 4, the structural part of the assertion holds $I = \beta \cup \gamma$.

It remains to show the algorithmic part of the assertion. In view of Lemma 4, one can construct γ in polynomial time. To determine I, we need to construct β in polynomial time. Start with $\beta = [n]$ and iteratively reduce β as follows. Using a polynomial-time algorithm for linear optimization, check if after a removal of one of the elements from β, the vector $\boldsymbol{v}$ is still in the conic hull of $\{\boldsymbol{a}_i : i \in \beta\}$. This procedure does at most n iterations. By Carathéodory's theorem, after the termination, the system of vectors $\boldsymbol{a}_i$ with $i \in \beta$ is linearly independent. □

Lemma 6 *Let $A \in \mathbb{Z}^{m \times n}$ be a matrix, whose columns positively span $\mathbb{R}^m$. Then $\mathcal{L}(A) = \mathrm{Sg}(A)$. Let $\boldsymbol{b} \in \mathcal{L}(A)$. If A and $\boldsymbol{b}$ are given in binary encoding, then a solution to $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$ can be constructed in polynomial time.*

Proof Since the columns of A positively span $\mathbb{R}^m$, the feasibility problem $A\boldsymbol{y} = \boldsymbol{0}$, $\boldsymbol{y} \in \mathbb{Q}^n_{\geq 1}$ of linear programming has a solution $\boldsymbol{y}$. One can determine such a solution $\boldsymbol{y}$ in polynomial time using a polynomial-time algorithm of linear optimization. The description size of $\boldsymbol{y}$ is polynomial. Thus, one can re-scale $\boldsymbol{y}$ to clear denominators in polynomial time, and arrive at a vector $\boldsymbol{y} \in \mathbb{Z}^n_{\geq 1}$ satisfying $A\boldsymbol{y} = \boldsymbol{0}$.

Now, if $\boldsymbol{b} \in \mathcal{L}(A)$. Then we first solve the Diophantine system $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n$ in polynomial time and determine one solution $\boldsymbol{x}^*$ of this system. But then the vector

$\boldsymbol{x} = \boldsymbol{x}^* + \|\boldsymbol{x}^*\|_\infty \boldsymbol{y}$ is a non-negative integer solution of $A\boldsymbol{x} = \boldsymbol{b}$. This verifies $\mathcal{L}(A) = \mathrm{Sg}(A)$ and shows the algorithmic part of the assertion. □

Proofs of Theorem 2 and Corollary 2 Choose $\tau \in \binom{[n]}{m}$ such that A_τ is non-singular and consdier I as in Lemma 5. In view of Lemma 6, one has $\mathcal{L}(A_I) = \mathrm{Sg}(A_I)$. Consequently, $\mathsf{ICR}(A) \leq \mathsf{ICC}(A) \leq |I|$ and the assertion of Theorem 2 follows from Lemma 5.

To show Corollary 2, observe that by Lemma 5, the set I can be constructed in polynomial time. To conclude the proof, it suffices to apply the algorithmic part of Lemma 6 to the sub-matrix A_I. □

4 Proofs of Theorems 3 and 4

To motivate Theorem 3 and its proof, we start this section by giving a self-containing proof of Theorem 4 which contains the key ideas to the proof of Theorem 3 while being less technical.

The following lemma is a key to the proof of Theorem 4.

Lemma 7 *Let $a_1, \ldots, a_t \in \mathbb{Z}_{>0}$, where $t \in \mathbb{Z}_{>0}$. If $t > 1 + \log_2(a_1)$, then the system*

$$\begin{aligned} &y_1 a_1 + \cdots + y_t a_t = 0, \\ &y_1 \in \mathbb{Z}_{\geq 0},\ y_2, \ldots, y_t \in \{-1, 0, 1\}. \end{aligned}$$

in the unknowns $y_1, \ldots, y_t$ has a solution that is not identically equal to zero.

Proof If X is a convex set whose affine hull has dimension at most k, then we use $\mathrm{vol}_k(X)$ to denote the k-dimensional volume of X.

Consider the convex set $Y \subseteq \mathbb{R}^t$ defined by $2t$ strict linear inequalities

$$\begin{aligned} -1 < &y_1 a_1 + \cdots + y_t a_t < 1, \\ -2 < &y_i < 2 \text{ for all } i \in \{2, \ldots, t\}. \end{aligned}$$

Clearly, the set Y is the interior of a hyper-parallelepiped and can also be described as $Y = \left\{\boldsymbol{y} \in \mathbb{R}^t : \|M\boldsymbol{y}\|_\infty < 1\right\}$, where M is the upper triangular matrix

$$M = \begin{pmatrix} a_1 & a_2 & \cdots & a_t \\ & 1/2 & & \\ & & \ddots & \\ & & & 1/2 \end{pmatrix}.$$

It is easy to see that $\mathrm{vol}_t(Y)$ of Y is

$$\mathrm{vol}_t(Y) = \mathrm{vol}_t(M^{-1}[-1, 1]^t) = \frac{1}{\det(M)} 2^t = \frac{4^t}{2a_1}.$$

The assumption $t > 1+\log_2(a_1)$ implies that the volume of Y is strictly larger than 2^t. Thus, by Minkowski's first theorem [6, Ch. VII, Sect. 3], the set Y contains a non-zero integer vector $\boldsymbol{y} = (y_1, \ldots, y_t)^\top \in \mathbb{Z}^t$. Without loss of generality we can assume that $y_1 \geq 0$ (if the latter is not true, one can replace $\boldsymbol{y}$ by $-\boldsymbol{y}$). The vector $\boldsymbol{y}$ is a desired solution from the assertion of the lemma. □

Proof of Theorem 4 By Lemma 3, we may assume that $\gcd(\boldsymbol{a}) = 1$. Further, without loss of generality, let $a_1 = \min\{a_1, \ldots, a_n\}$. Let $b \in \mathrm{Sg}(\boldsymbol{a})$. By the definition of $\mathsf{ICR}(\boldsymbol{a})$, the integer Carathéodory rank, we need to show the existence of solution to $\boldsymbol{ax} = b$, $\boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$ satisfying $\|\boldsymbol{x}\|_0 \leq 1 + \log_2(a_1)$. Choose a solution $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ with the property that the number of indices $i \in \{2, \ldots, n\}$ for which $x_i \neq 0$ is minimized. Without loss of generality, we can assume that, for some $t \in \{2, \ldots, n\}$ one has $x_2 > 0, \ldots, x_t > 0, x_{t+1} = \cdots = x_n = 0$. We claim that Lemma 7 implies $t \leq 1 + \log_2(a_1)$. In fact, if the latter was not true, then a solution $\boldsymbol{y} \in \mathbb{R}^t$ of the system in Lemma 7 could be extended to a solution $\boldsymbol{y} \in \mathbb{R}^n$ by appending zero components. It is clear that some of the components $y_2, \ldots, y_t$ are negative, because $a_2 > 0, \ldots, a_t > 0$. It then turns out that, for an appropriate choice of $k \in \mathbb{Z}_{\geq 0}$, the vector $\boldsymbol{x}' = (x'_1, \ldots, x'_n)^\top = \boldsymbol{x} + k\boldsymbol{y}$ satisfies $\boldsymbol{ax} = b$ and $x'_1 \geq 0, \ldots, x'_t \geq 0$, $x'_{t+1} = \cdots = x'_n = 0$ and $x'_i = 0$ for at least one $i \in \{2, \ldots, t\}$. Indeed, one can choose k to be the minimum among all x_i with $i \in \{2, \ldots, t\}$ and $y_i = -1$.

The existence of $\boldsymbol{x}'$ with at most $t-1$ non-zero components x'_i with $i \in \{2, \ldots, n\}$ contradicts the choice of $\boldsymbol{x}$ and yields the assertion. □

To prove Theorem 3 we need two auxiliary lemmas, of which one generalizes Lemma 7. In what follows, we will denote the linear hull of $X \subseteq \mathbb{R}^m$ by $\mathrm{lin}(X)$.

Lemma 8 *Let $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{Z}^{m\times n}$ satisfy (6)–(8). Then there exists a basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ of the lattice $\mathcal{L}(A)$ satisfying*

$$\mathrm{cone}(\boldsymbol{a}_1) = \mathrm{cone}(\boldsymbol{u}_1) \quad \textit{and} \quad \mathrm{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)\backslash\{0\} \subseteq H,$$

where H is the open halfspace

$$H := \left\{ \sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i \ : \ \alpha_1 > 0, \alpha_2, \ldots, \alpha_m \in \mathbb{R} \right\}.$$

Proof We argue by induction on the dimension m. For $m = 1$, the assertion holds with $u_1 = \gcd(A)$. Assume the assertion is true in dimension $m-1$, where $m \geq 2$. In view (7) and (8), $\mathrm{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ is an m-dimensional pointed cone and so it has a facet F that contains $\mathrm{cone}(\boldsymbol{a}_1)$ as a subset. The semigroup $\mathcal{L}(A) \cap F$ is finitely generated so that one can choose finitely many vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t$ satisfying $\mathcal{L}(A) \cap F = \mathcal{S}\}(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t)$. Since $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t$ contain generators of the extreme rays of F and $\boldsymbol{a}_1$ generates one of these extreme rays, we can assume that $\mathrm{cone}(\boldsymbol{b}_1) = \mathrm{cone}(\boldsymbol{a}_1)$. Viewing $\mathrm{lin}(F)$ as $\mathbb{R}^{m-1}$ and applying the induction assumption to the cone $F = \mathrm{cone}(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t)$, we conclude that there exists a basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{m-1} \in \mathbb{Z}^m$ of

the lattice $\mathcal{L}(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t) = \mathcal{L}(A) \cap \operatorname{lin}(F)$ that satisfies the conditions $\operatorname{cone}(\boldsymbol{a}_1) = \operatorname{cone}(\boldsymbol{b}_1) = \operatorname{cone}(\boldsymbol{u}_1)$ and

$$F \subseteq \{0\} \cup \left\{ \sum_{i=1}^{m-1} \alpha_i \boldsymbol{u}_i \ : \ \alpha_1 > 0, \alpha_2, \ldots, \alpha_{m-1} \in \mathbb{R} \right\}.$$

To conclude the proof, we need to choose $\boldsymbol{u}_m$ appropriately. We extend the basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{m-1}$ of $\mathcal{L}(A) \cap \operatorname{lin}(F)$ to a basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{m-1}, \boldsymbol{v}$ of $\mathcal{L}(A)$. We will fix

$$\boldsymbol{u}_m := \boldsymbol{v} - N(\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_{m-1})$$

with an appropriate $N \in \mathbb{Z}_{>0}$.

Since $\operatorname{lin}(F)$ is a supporting hyperplane of $\operatorname{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$, we can assume, after possibly exchanging the roles of $\boldsymbol{v}$ and $-\boldsymbol{v}$, that $\boldsymbol{v}$ and $\operatorname{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ are on the same side of the hyperplane $\operatorname{lin}(F)$. Every vector $\boldsymbol{a}_i$ can be expressed as

$$\boldsymbol{a}_i = \sum_{j=1}^{m-1} \beta_{i,j} \boldsymbol{u}_j + \beta_i \boldsymbol{v},$$

where $\beta_{i,j} \in \mathbb{Z}$ and $\beta_i \in \mathbb{Z}_{\geq 0}$. Furthermore, if $\boldsymbol{a}_i \notin F$, then, in view of (6), $\beta_i > 0$ and, thus, using the representation

$$\boldsymbol{a}_i = \sum_{j=1}^{m-1} (\beta_{i,j} + N\beta_i) \boldsymbol{u}_j + \beta_i \underbrace{(\boldsymbol{v} - N(\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_{m-1}))}_{\boldsymbol{u}_m}$$

we see that, for N large enough, one has $\beta_{i,j} + N\beta_i > 0$ so that $\boldsymbol{a}_i \in H$ for all vectors $\boldsymbol{a}_i$ that do not belong to F. The vectors $\boldsymbol{a}_i$ that belong to F are in H by the choice of $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{m-1}$.

Consequently, $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} \subseteq H$. Since $\operatorname{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ is pointed, the latter implies $\operatorname{cone}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \backslash \{\boldsymbol{0}\} \subseteq H$. □

The following lemma generalizes Lemma 7.

Lemma 9 *Let $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{Z}^{m \times n}$ satisfy (6)–(8). Let*

$$n > m + \log_2 \left(\frac{q(A)}{\gcd(A)} \right), \tag{13}$$

where $q(A) := \sqrt{\sum_{I \in \binom{[n]}{m} : 1 \in I} \det(A_I)^2}$. Then the system

$$\begin{aligned} &y_1 \boldsymbol{a}_1 + \cdots + y_n \boldsymbol{a}_n = \boldsymbol{0}, \\ &y_1 \in \mathbb{Z}_{\geq 0}, \ y_2, \ldots, y_n \in \{-1, 0, 1\}. \end{aligned}$$

in the unknowns $y_1, \dots, y_n$ *has a solution such that at least one of* $y_2, \dots, y_n$ *equals* -1.

Proof In view of Lemma 7 we may assume that $m \geq 2$. Consider $U = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m) \in \mathbb{Z}^{m\times m}$, where $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$ are vectors as in Lemma 8. We can express the columns of A in the basis $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$. This means that $A = U\overline{A}$ holds for some matrix $\overline{A} = (\overline{a}_{i,j})_{i\in[m],\ j\in[n]} \in \mathbb{Z}^{m\times n}$. In view of Lemma 8, we have $\overline{a}_{1,j} > 0$ for every $j \in [n]$ and $\overline{a}_{i,1} = 0$ for every $i \in \{2, \dots, m\}$. Clearly, $A\boldsymbol{y} = 0$ is equivalent to $\overline{A}\boldsymbol{y} = 0$. The system $\overline{A}\boldsymbol{y} = 0$ has the structure

$$\underbrace{\left(\begin{array}{c|ccc} \overline{a}_{1,1} & \overline{a}_{1,2} & \cdots & \overline{a}_{1,n} \\ \hline 0 & \overline{a}_{2,2} & \cdots & \overline{a}_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & \overline{a}_{m,2} & \cdots & \overline{a}_{m,n} \end{array}\right)}_{\overline{A}} \cdot \underbrace{\begin{pmatrix} y_1 \\ \hline y_2 \\ \vdots \\ y_n \end{pmatrix}}_{y} = \begin{pmatrix} 0 \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{14}$$

Using the notation $\alpha_1, \dots, \alpha_n \in \mathbb{R}_{>0}$ and $B \in \mathbb{Z}^{(m-1)\times(n-1)}$ such that

$$\overline{A} = \left(\begin{array}{c|ccc} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right), \tag{15}$$

we formulate (14) as

$$\alpha_1 y_1 + \cdots + \alpha_n y_n = 0, \qquad \begin{pmatrix} y_2 \\ \vdots \\ y_n \end{pmatrix} \in \ker(B). \tag{16}$$

By the m-dimensionality assumption in (7), the matrix A has rank m. In view of this and $\alpha_1 > 0$, the matrix B has rank $m-1$. Therefore

$$\Lambda := \mathbb{Z}^{n-1} \cap \ker(B)$$

is a $(n-m)$–dimensional lattice with

$$\det(\Lambda) = \frac{\sqrt{\det(BB^\top)}}{\gcd(B)}$$

(recall that we assume $m \geq 2$ and hence Λ is well-defined).

It turns out that $\gcd(B) = 1$. Indeed, by the choice of $\overline{A}$, we have $\mathcal{L}(\overline{A}) = \mathbb{Z}^m$. By (15), the lattice $\mathcal{L}(B)$ is obtained by projecting $\mathcal{L}(\overline{A})$ onto the last $m-1$ components. So, $\mathcal{L}(B) = \mathbb{Z}^{m-1}$ and by this $\gcd(B) = 1$. This yields

$$\det(\Lambda) = \sqrt{\det(BB^\top)}.$$

We introduce the $\mathbf{0}$-symmetric convex set

$$C := (-2, 2)^{n-1} \cap \ker(B),$$

which is the relative interior of an $(n - m)$-dimensional cross-section of the cube $[-2, 2]^{n-1}$. By Vaaler's cube slicing inequality [33], we have

$$\mathrm{vol}_{n-m}(C) \geq 4^{n-m}.$$

We introduce the $\mathbf{0}$-symmetric convex set

$$C' = \left\{ (y_1, \ldots, y_n)^\top \,:\, -1 < \alpha_1 y_1 + \cdots + \alpha_n y_n < 1,\ (y_2, \ldots, y_n)^\top \in C \right\}.$$

It is easy to see that

$$\mathrm{vol}_{n-m+1}(C') = \frac{2\,\mathrm{vol}_{n-m}(C)}{\alpha_1}.$$

By Minkowski's first theorem applied to the set C' and the lattice $\mathbb{Z} \times \Lambda$, we know that if

$$\mathrm{vol}_{n-m+1}(C') > 2^{n-m+1} \det(\mathbb{Z} \times \Lambda),$$

then the set C' contains a non-zero vector $\boldsymbol{y}$ of the lattice $\mathbb{Z} \times \Lambda$. We have

$$\det(\mathbb{Z} \times \Lambda) = \det(\Lambda) = \sqrt{\det(BB^\top)}$$

and

$$\mathrm{vol}_{n-m+1}(C') = \frac{2\,\mathrm{vol}_{n-m}(C)}{\alpha_1} \geq \frac{2 \cdot 4^{n-m}}{\alpha_1}.$$

This means, that the assumptions of Minkowski's first theorem are fulfilled when

$$\frac{2 \cdot 4^{n-m}}{\alpha_1} > 2^{n-m+1} \sqrt{\det(BB^\top)}. \tag{17}$$

Clearly, (17) can be written as

$$n > m + \log_2 \left(\alpha_1 \sqrt{\det(BB^\top)} \right). \tag{18}$$

We show that (13) and (18) are equivalent by verifying that the terms under the logarithm coincide. Keeping in mind that B is a submatrix of $\overline{A}$ [see (15)], we index

columns of B by $\{2, \dots, n\}$. We have

$$\alpha_1 \sqrt{\det(BB^\top)} = \alpha_1 \sqrt{\sum_{J \in \binom{\{2,\dots,n\}}{m-1}} \det(B_J)^2}$$

$$= \sqrt{\sum_{J \in \binom{\{2,\dots,n\}}{m-1}} (\alpha_1 \det(B_J))^2} = \sqrt{\sum_{J \in \binom{\{2,\dots,n\}}{m-1}} \det(\bar{A}_{\{1\} \cup J})^2}$$

$$= q(\bar{A}) = q(U^{-1}A) = \frac{q(A)}{|\det(U)|} = \frac{q(A)}{\gcd(A)},$$

which verifies the equivalence of (13) and (18).

Now, consider a non-zero lattice vector $\boldsymbol{y} = (y_1, \dots, y_n)^\top$ in $C' \cap (\mathbb{Z} \times \Lambda)$. By the choice of C' and Λ, the vector $\boldsymbol{y}$ is a solution of $\bar{A}\boldsymbol{y} = 0$ and by this also a solution of $A\boldsymbol{y} = 0$ and, furthermore, we have $y_2, \dots, y_n \in \{-1, 0, 1\}$. Possibly replacing $\boldsymbol{y}$ with $-\boldsymbol{y}$, we can ensure that $y_1 \geq 0$. Since the equation $\alpha_1 y_1 + \dots + \alpha_n y_n = 0$ (contained in the system $\bar{A}\boldsymbol{y} = 0$) has positive coefficients and since $y_1 \geq 0$, we conclude that at least one of the variables $y_2, \dots, y_n$ of our solution $\boldsymbol{y}$ is negative. Thus, our solution $\boldsymbol{y}$ satisfies the assertions of the lemma. □

Proof of Theorem 3 It is sufficient to show that any feasible problem $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$ with the matrix A satisfying assumptions (6)–(8) has a solution with the size of support bounded by $m + \log_2(q(A)/\gcd(A))$. Choose a solution $\boldsymbol{x}^* = (x_1^*, \dots, x_n^*)^\top$ of $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$, for which the number of indices $i \in \{2, \dots, n\}$ satisfying $x_i^* \neq 0$ is minimal. After a reordering the columns of the matrix A, we can assume that $x_i^* \neq 0$ for $2 \leq i \leq s$ and $x_i^* = 0$ for $i > s$, for some choice of s. Consider the vector space $V := \mathrm{lin}(\boldsymbol{a}_1, \dots, \boldsymbol{a}_s)$. Since $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ linearly span $\mathbb{R}^m$, among the vectors $\boldsymbol{a}_{s+1}, \dots, \boldsymbol{a}_n$ one can choose linearly independent columns that together with a basis of V form a basis of $\mathbb{R}^m$. Without loss of generality, let $\boldsymbol{a}_{s+1}, \dots, \boldsymbol{a}_t$ be such vectors, that is, one has $V \oplus W = \mathbb{R}^m$ with $W := \mathrm{lin}(\boldsymbol{a}_{s+1}, \dots, \boldsymbol{a}_t)$. In the degenerate case $V = \mathbb{R}^m$, we just fix $t = s$ and $W = \{\boldsymbol{0}\}$. We show that

$$t \leq m + \log_2\left(\frac{q(A)}{\gcd(A)}\right)$$

arguing by contradiction. Assume that $t > m + \log_2\left(\frac{q(A)}{\gcd(A)}\right)$. Conditions (6)–(8) are fulfilled for the matrix $A_{[t]} = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_t)$ in place of $A = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$. Since $q(A_{[t]}) \leq q(A)$ and $\gcd(A_{[t]}) \geq \gcd(A)$, we have

$$t > m + \log_2\left(\frac{q(A_{[t]})}{\gcd(A_{[t]})}\right).$$

The application of Lemma 9 to the matrix $A_{[t]}$ yields the existence of a vector $\boldsymbol{y} = (y_1, \dots, y_t)^\top \in \mathbb{Z}_{\geq 0} \times \{-1, 0, 1\}^{t-1}$ satisfying $A_{[t]}\boldsymbol{y} = \boldsymbol{0}$ with at least one of the

values $y_2, \ldots, y_t$ equal to -1. In view of $V \oplus W = \mathbb{R}^m$ and the linear independence of the vectors $\boldsymbol{a}_{s+1}, \ldots, \boldsymbol{a}_t$, the equality

$$\mathbf{0} = A_{[t]}\boldsymbol{y} = \sum_{i=1}^{t} \boldsymbol{a}_i y_i = \underbrace{\sum_{i=1}^{s} \boldsymbol{a}_i y_i}_{\in V} + \underbrace{\sum_{i=s+1}^{t} \boldsymbol{a}_i y_i}_{\in W}$$

yields $y_i = 0$ for $s < i \leq t$. This shows that one of the values $y_2, \ldots, y_s$ is equal to -1. We convert $\boldsymbol{y} \in \mathbb{Z}^t$ to a vector in $\boldsymbol{y}' \in \mathbb{Z}^n$ by appending zero components.

Clearly, $\boldsymbol{x} = \boldsymbol{x}^* + k\boldsymbol{y}'$ is a solution of $A\boldsymbol{x} = \boldsymbol{b}$, and if we choose k to be the minimum among the values x_i^*, where $i \in \{2, \ldots, s\}$ and $y_i = -1$, then $\boldsymbol{x} = \boldsymbol{x}^* + k\boldsymbol{y}'$ is a solution of $A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$, for which the number of indices $i \in \{2, \ldots, n\}$ satisfying $x_i^* \neq 0$ is smaller than for the solution $\boldsymbol{x}^*$ of $A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$. This contradicts the choice of $\boldsymbol{x}^*$ and shows that

$$\|\boldsymbol{x}^*\|_0 \leq t \leq m + \log_2\left(\frac{q(A)}{\gcd(A)}\right). \qquad \square$$

5 Optimality of the bounds

In this section we prove a series of three propositions that show, respectively, that the Theorems 1, 2 and 3 are optimal. For this we introduce the following notation. For an integer $z \in \mathbb{Z}_{>0}$ with prime factorization $z = p_1^{s_1} \cdots p_k^{s_k}$, where the distinct prime factors $p_1, \ldots, p_k$ have the multiplicities $s_1, \ldots, s_k \in \mathbb{Z}_{>0}$, we define the set

$$S(z) := \left\{ \frac{z}{p_i^{s_i}} \ : \ i \in [k] \right\}.$$

The elements of $S(z)$ are relatively prime, but every non-empty proper subset of $S(z)$ has a common divisor larger than one. The set $S(z)$ has $\omega(z)$ elements. If z is a prime number, we have $S(z) = \{1\}$.

Proposition 1 *Let $m \in \mathbb{Z}_{>0}$ and let $F : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a function providing the bound*

$$\mathit{ILR}(A) \leq \min\left\{ F\left(\frac{|\Delta|}{\gcd(A)}\right) \ : \ \Delta \text{ nonzero } m \times m \text{ minor of } A \right\}$$

for all $n \in \mathbb{Z}_{\geq m}$ and all matrices $A \in \mathbb{Z}^{m \times n}$ of full row rank. Then $F(z) \geq m + \Omega_m(z)$ holds for every $z \in \mathbb{Z}_{>0}$.

Proof Let $z \in \mathbb{Z}_{>0}$. We need to show $F(z) \geq m + \Omega_m(z)$. For $z = 1$, this reduces to showing $F(z) \geq m$. The latter is clear, because the matrices A in the formulation of the assertion have rank m. We now consider the case $z \geq 2$. We decompose z as $z = z_1 \cdots z_m$ into m factors $z_1, \ldots, z_m \in \mathbb{Z}_{>0}$ as follows. Let α_p denote the multiplicity

of the prime number p in the prime factorization of z, i.e., $z = \prod_{p \text{ prime}} p^{\alpha_p}$. Then we define for $i = 1, \ldots, m-1$

$$z_i := \prod_{\substack{p \text{ prime} \\ \alpha_p - i \geq 0}} p \quad \text{and} \quad z_m := \prod_{\substack{p \text{ prime} \\ \alpha_p - m \geq 0}} p^{\alpha_p - m + 1}.$$

Let q be prime such that $\gcd(z, q) = 1$. Consider the sets $S(z_1 q), \ldots, S(z_m q)$. Note that, by construction, $z_i \in S(z_i q)$ for every $i \in [m]$. Finally, we fix A to be a matrix of size $m \times n$ with $n = \sum_{i=1}^{m} |S(z_i q)|$ and the columns $r\boldsymbol{e}_i$, where $i \in [m]$ and $r \in S(z_i q)$. (The order of these columns in A does not matter.) By construction, $\gcd(A) = 1$, $\mathsf{ILR}(A) = n$ and $\Delta = \pm z_1 \cdots z_m = \pm z$ is an $m \times m$ minor of A. Consequently, $n \leq F(z)$. It remains to express n in terms of z. We have $|S(z_i q)| = \omega(z_i q) = \omega(z_i) + 1$. By construction, each prime factor of z of multiplicity $\mu \leq m$ has contributed one unit into μ of the values $\omega(z_1), \ldots, \omega(z_m)$ and each prime factor of z of multiplicity $\mu \geq m$ has contributed one unit to all m values $\omega(z_1), \ldots, \omega(z_m)$. This shows $\sum_{i=1}^{m} \omega(z_i) = \Omega_m(z)$ and implies $F(z) \geq n = m + \Omega_m(z)$. □

Proposition 2 *Let $m \in \mathbb{Z}_{>0}$ and let $F : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a function providing the bound*

$$\mathit{ICR}(A) \leq \min \left\{ F\left(\frac{|\Delta|}{\gcd(A)} \right) : \Delta \text{ nonzero } m \times m \text{ minor of } A \right\}$$

for all $n \in \mathbb{Z}_{>0}$ and all matrices $A \in \mathbb{Z}^{m \times n}$ whose columns positively span $\mathbb{R}^m$. Then $F(z) \geq 2m + \Omega_m(z)$ holds for every $z \in \mathbb{Z}_{>0}$.

Proof To deal with $z = 1$, just fix A to be the matrix with $2m$ columns $2\boldsymbol{e}_i$ and $-\boldsymbol{e}_i$, where $i \in [m]$. For $z > 1$ adapt the proof idea of Proposition 1. Decompose z into the product $z = z_1 \cdots z_m$ as in the proof of Proposition 1, choose a prime q that is not a factor of z and introduce A to be a matrix with the columns $r\boldsymbol{e}_i$, where $i \in [m]$ and $r \in S(z_i q) \cup \{-z_i q\}$. The matrix A has $2m + \Omega_m(z)$ columns, and it satisfies $\gcd(A) = 1$ and $\mathsf{ICR}(A) \geq 2m + \Omega_m(z)$. On the other hand, as in the proof of Proposition 1, one can see that $\pm z$ is an $m \times m$ minor of A. All this shows $F(z) \geq 2m + \Omega_m(z)$. □

Proposition 3 *Let $m \in \mathbb{Z}_{>0}$ and let $F : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a non-decreasing function that yields the bound*

$$\mathit{ICR}(A) \leq F\left(\frac{q(A)}{\gcd(A)} \right)$$

for all $n \in \mathbb{Z}_{\geq m}$ and all matrices $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{Z}^{m \times n}$ that satisfy conditions (6)–(8). Then $F(z) \geq m + \lfloor \log_2(z) \rfloor$ holds for every $z \in \mathbb{Z}_{>0}$.

Proof In view of the monotonicity of F, it suffices to consider $z = 2^s$ with $s \in \mathbb{Z}_{\geq 0}$. For $s = 0$, choose A to be the identity matrix. For $s \geq 1$, we set $n = s + m$ and define $A = (\alpha_1 \boldsymbol{e}_1, \ldots, \alpha_{s+1} \boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m)$, where $\alpha_1 = 2^s$, $\alpha_i = 2^{s-1+i} + 2^{s+1-i}$ for

$i \in \{2, \ldots, s+1\}$. One has $\gcd(A) = 1$, which can be verified directly for $s \leq 2$ and follows from $\gcd(\alpha_1, \alpha_2, \alpha_3, \alpha_{s+1}) = \gcd(2^s, 2^{s-1} \cdot 5, 2^{s-2} \cdot 17, 2^{2s} + 1) = 1$ for $s \geq 3$. Further, $q(A) = \alpha_1 = 2^s$. To show $\mathsf{ICR}(A) = m + s$, let $\boldsymbol{b} = (2^{2s+1} - 1, 1, \ldots, 1)$. It is clear that $\boldsymbol{x} = (1, \ldots, 1)^\top$ is a solution, and it remains to show that this is the only solution. The number $2^{2s+1} - 1$ is odd, while α_s is the only odd coefficient on the left-hand side of the first equation. Thus, for every solution, one has $x_s \geq 1$. Since $2a_s > 2^{2s+1} - 1$, we even have $x_s = 1$. Substituting $x_s = 1$, and dividing the first equation by 2, we arrive at the same system of equations with $m + s - 1$ variables. Iterating we conclude that all variables $x_1, \ldots, x_{m+s}$ must be equal to 1. This shows $\mathsf{ICR}(A) = m + s$ and yields the desired inequality $F(z) \geq m + s$. □

6 Results on computational complexity

In this final section, we explore the computational complexity of the functions $\mathsf{ILR}(A)$, $\mathsf{ILC}(A)$, $\mathsf{ICR}(A)$ and $\mathsf{ICC}(A)$. We begin with the hardness results of Theorem 5.

6.1 Proof of Theorem 5, Parts (i) and (ii)

In the case $m = 1$ of just one row, we use the notation $\boldsymbol{a} = A$ and denote by a_i the i-th component of A. As before, we use the notation $\boldsymbol{a}_\tau$ for $\tau \subseteq [n]$. It turns out that in the case of one row, two of our four functions coincide:

Proposition 4 *For all $\boldsymbol{a} \in \mathbb{Z}^{1\times n}$, one has $\mathsf{ILR}(\boldsymbol{a}) = \mathsf{ILC}(\boldsymbol{a})$.*

Proof The inequality $\mathsf{ILR}(\boldsymbol{a}) \leq \mathsf{ILC}(\boldsymbol{a})$ is clear from the definition. We show the reverse inequality $\mathsf{ILC}(\boldsymbol{a}) \leq \mathsf{ILR}(\boldsymbol{a})$. Assume $\boldsymbol{a} \neq \boldsymbol{0}$, since otherwise the assertion is trivial. Since $\boldsymbol{a}$ can be divided by $\gcd(\boldsymbol{a})$, there is no loss of generality in assuming $\gcd(\boldsymbol{a}) = 1$. Let $k := \mathsf{ILR}(\boldsymbol{a})$. Taking into account the definition of ILR and $\mathcal{L}(\boldsymbol{a}) = \gcd(\boldsymbol{a})\mathbb{Z} = \mathbb{Z}$, we obtain

$$\mathbb{Z} = \bigcup_{\tau \in \binom{[n]}{k}} \mathcal{L}(\boldsymbol{a}_\tau) = \bigcup_{\tau \in \binom{[n]}{k}} \gcd(\boldsymbol{a}_\tau)\mathbb{Z}. \tag{19}$$

We claim that $\gcd(\boldsymbol{a}_\tau) = 1$ holds for some $\tau \in \binom{[n]}{k}$. Indeed, if $\gcd(\boldsymbol{a}_\tau) \geq 2$ for all $\tau \in \binom{[n]}{k}$, then the number $z := 1 + \prod_{\tau \in \binom{[n]}{k}} \gcd(\boldsymbol{a}_\tau)$, which is relatively prime with each $\gcd(\boldsymbol{a}_\tau)$, does not belong to any of the sets $\gcd(\boldsymbol{a}_\tau)\mathbb{Z}$ with $\tau \in \binom{[n]}{k}$. This is a contradiction to (19). Thus, $\mathsf{ILC}(\boldsymbol{a}) \leq k = \mathsf{ILR}(\boldsymbol{a})$. □

The following result shows how to reduce ILC to ICR and ICC.

Proposition 5 *Let $\boldsymbol{a} \in \mathbb{Z}^{1\times n}_{\geq 2}$ with $\gcd(\boldsymbol{a}) = 1$ and let $\pi := \prod_{i=1}^n a_i$ and $\boldsymbol{a}^+ := (\boldsymbol{a}, -\pi) \in \mathbb{Z}^{1\times(n+1)}$. Then $\mathsf{ICR}(\boldsymbol{a}^+) = \mathsf{ICC}(\boldsymbol{a}^+) = 1 + \mathsf{ILC}(\boldsymbol{a})$.*

Proof It suffices to check the validity of the inequalities

$$\mathsf{ICR}(\boldsymbol{a}^+) \leq \mathsf{ICC}(\boldsymbol{a}^+) \leq \mathsf{ILC}(\boldsymbol{a}) + 1 \leq \mathsf{ICR}(\boldsymbol{a}^+).$$

The first inequality follows directly from the definitions of ICR and ICC. Let us show $\mathsf{ICC}(\boldsymbol{a}^+) \leq 1+k$ for $k := \mathsf{ILC}(\boldsymbol{a})$. Consider a k-element set $\tau \subseteq [n]$ such that $\gcd(\boldsymbol{a}_\tau) = 1$. Without loss of generality, let $\tau = [k]$ so that $\gcd(a_1, \ldots, a_k) = 1$. Then there exists $z_1, \ldots, z_k \in \mathbb{Z}$ such that $1 = \sum_{i=1}^{k} z_i a_i$. We now simultaneously show $\mathrm{Sg}(\boldsymbol{a}^+) = \mathbb{Z}$ and $\mathsf{ICC}(\boldsymbol{a}^+) \leq k+1$. One clearly has $y_{n+1}(-\pi) + \sum_{i=1}^{k} y_i a_i = 0$ with $y_i := \frac{\pi}{a_i}$ for $i \in [k]$ and $y_{n+1} = k$. Consequently, for every $b \in \mathbb{Z}$, the equality $b = N y_{n+1} a_{n+1} + \sum_{i=1}^{k} (N y_i + b z_i) a_i$ holds for an arbitrary $N \in \mathbb{Z}_{>0}$. If we choose N large enough, all of the coefficients in the above representation become non-negative. This shows that every $b \in \mathbb{Z}$ belongs to $\mathrm{Sg}(\boldsymbol{a}^+)$ and, since we have used $k+1$ generators from $\boldsymbol{a}^+$ to represent b, we have $\mathsf{ICC}(\boldsymbol{a}^+) \leq k+1$.

To conclude the proof, it remains to verify $k+1 \leq \mathsf{ICR}(\boldsymbol{a}^+)$. It is easy to check that 1 is an element of the semigroup $\mathrm{Sg}(\boldsymbol{a}^+) = \mathbb{Z}$ that cannot be represented using at most k of the $n+1$ generators $a_1, \ldots, a_n, -\pi$. Indeed, the only negative generator $-\pi$ has to be used. If, apart from this generator, one uses at most $k-1$ positive generators, by the definition of k, the chosen generators have the gcd strictly larger than one, which is a contradiction. □

We will also make use of the hardness of the *set-cover problem*, which is the following classical NP-complete problem, see [19, Problem: SP5]. The input of the set-cover problem consists of $k, t \in \mathbb{Z}_{>0}$ and a family $\mathcal{S} := \{S_1, \ldots, S_n\}$ of n sets with $S_1 \cup \cdots \cup S_n = [t]$. We use $\mathrm{mincov}(\mathcal{S})$ to denote the minimal cardinality of $\tau \subseteq [n]$ such that $\bigcup_{i \in \tau} S_i = [t]$ holds. The set-cover problem is the problem to decide whether $\mathrm{mincov}(\mathcal{S}) \leq k$ holds.

Proof of Theorem 5, Parts (i) and (ii) Since $m = 1$, we use the notation $\boldsymbol{a} := A$. By Proposition 4, it is sufficient to consider only the three decision problems $\mathsf{ILC}(\boldsymbol{a}) \leq k$, $\mathsf{ICC}(\boldsymbol{a}) \leq k$ and $\mathsf{ICR}(\boldsymbol{a}) \leq k$.

We assume $\boldsymbol{a} \neq \boldsymbol{0}$. For $\tau \subseteq [n]$, one has $\mathcal{L}(\boldsymbol{a}) = \gcd(\boldsymbol{a}_\tau)\mathbb{Z}$. Hence, $\mathsf{ILC}(\boldsymbol{a})$ is the minimum cardinality of a set τ that satisfies $\gcd(\boldsymbol{a}) = \gcd(\boldsymbol{a}_\tau)$. Thus, the validity of $\mathsf{ILC}(\boldsymbol{a}) \leq k$ is certified by a set $\tau \subseteq [n]$ with at most k elements for which $\gcd(\boldsymbol{a}) = \gcd(\boldsymbol{a}_\tau)$ is fulfilled. Since the gcd is computable in polynomial time, this shows that our decision problem is in NP. In order to show that deciding $\mathsf{ICC}(\boldsymbol{a}) \leq k$ is in NP, we can use a certificate consisting of a set τ with $\mathsf{ICC}(\boldsymbol{a}_\tau) \leq k$ and solutions of problems $\boldsymbol{a}_\tau \boldsymbol{x} = a_i$, $\boldsymbol{x} \in \mathbb{Z}_{\geq 0}^{|\tau|}$, with $i \in [n]$, that have a polynomial description size.

To prove hardness of $\mathsf{ILC}(\boldsymbol{a}) \leq k$, we will use reduction from the *set-cover problem*. Consider a family $\mathcal{S} = \{S_1, \ldots, S_n\}$ with $\bigcup_{i=1}^{n} S_i = [t]$. Since every of the t elements from $[t]$ occurs in some of the sets $S_1, \ldots, S_n$ and since we have n sets in total, it is clear that the size for the input the set-cover problem is of order *at least* $n+t$. The reduction is as follows. We compute the first t prime numbers $p_1, \ldots, p_t$. To this end, we can use a weaker version of the Prime Number Theorem, established by Chebyshev, which asserts that for $t \geq 2$ there exists a universal constant $c > 0$, such that $p_t \leq c\, t \log_e t$, see [23, Theorem 9]. Hence, $p_1, \ldots, p_t$ can be found by running the sieve of Eratosthenes or some more brute-force algorithm on the range of integers $\{1, \ldots, O(t \log_e t)\}$.

We are going to encode elements of $\{1, \ldots, t\}$ via the above prime numbers. Accordingly, we encode sets $S_1, \ldots, S_n$ via integer numbers as follows: with S_j we associate

$a_j := \prod_{i \in [t] \setminus S_j} p_i$. This means that a_j is the product of those prime numbers p_i whose index i is not in S_j. As the prime numbers $p_1, \ldots, p_t$ have a polynomial bit size in t, the numbers $a_1, \ldots, a_n$ can be computed in polynomial time.

Since the union of $S_1, \ldots, S_n$ is $[t]$, we conclude that $\gcd(\boldsymbol{a}) = 1$ holds for $\boldsymbol{a} = (a_1, \ldots, a_n) \in \mathbb{Z}^{1\times n}$. More generally, $\bigcup_{j\in\tau} S_i = [t]$ holds if and only if $\gcd(\boldsymbol{a}_\tau) = 1$. This shows $\mathrm{mincov}(\mathcal{S}) = \mathsf{ILC}(\boldsymbol{a})$. Thus, polynomial-time reduction $\mathcal{S} \mapsto \boldsymbol{a}$ converts the set-cover problem $\mathrm{mincov}(\mathcal{S}) \le k$ to the problem $\mathsf{ILC}(\boldsymbol{a}) \le k$.

In view of Proposition 5, the computation of $\mathsf{ILC}(\boldsymbol{a})$ for $\boldsymbol{a} \in \mathbb{Z}^{1\times n}_{\ge 2}$ satisfying $\gcd(\boldsymbol{a}) = 1$ can be reduced, in polynomial time, to computation of ICR and ICC in the case of one row, by constructing the vector $\boldsymbol{a}^+$ out of $\boldsymbol{a}$. Thus, the NP-hardness of deciding $\mathsf{ICR}(\boldsymbol{a}) \le k$ and $\mathsf{ICC}(\boldsymbol{a}) \le k$ follows form the NP-hardness of deciding $\mathsf{ILC}(\boldsymbol{a}) \le k$. □

The exact complexity status for the analogous decision problem $\mathsf{ICR}(\boldsymbol{a}) \le k$ remains unresolved. It is neither clear if the latter decision problem is in NP nor is it clear if this problem is in co-NP.

6.2 Proof of Theorem 5, Parts (iii) and (iv)

We recall that a computational problem is called *strongly NP-hard* if it is NP-hard with respect to the unary encoding of the coefficients of the input. Applied to our setting, this means that the coefficients of $A \in \mathbb{Z}^{m\times n}$ are given in the unary encoding. A decision problem is called *strongly NP-complete* if it belongs to NP (with respect to the binary encoding of the coefficients) and is strongly NP-hard.

For a set $S \subseteq [m]$, let $\chi_S \in \{0, 1\}^m$ be the characteristic vector of S.

Lemma 10 *Let $m \in \mathbb{Z}_{>0}$ and let $\mathcal{S} = \{S_1, \ldots, S_\ell\}$ be a family of sets with $\bigcup_{i=1}^{\ell} S_i = [m]$. Then, for the matrix*

$$A := (-p_1\boldsymbol{e}_1, \ldots, -p_m\boldsymbol{e}_m, q\chi_{S_1}, \ldots, q\chi_{S_m}) \in \mathbb{Z}^{m\times(m+\ell)},$$

defined using $\mathcal{S}$ and arbitrary $m+1$ pairwise distinct prime numbers $p_1, \ldots, p_m$ and q, one has $\mathsf{ILC}(A) = \mathsf{ILR}(A) = \mathsf{ICC}(A) = \mathsf{ICR}(A) = \mathrm{mincov}(\mathcal{S}) + m$.

Proof We first show $\gcd(A) = 1$. The minor from the first m columns of A is equal to $p_1 \cdots p_m$ in the absolute value. For $i \in [m]$, the minor obtained by taking the first m columns, except for the i-th one, and a column $q\chi_{S_j}$ with a set S_j satisfying $i \in S_j$, is equal to $p_1 \cdots p_{i-1}p_{i+1} \cdots p_m q$ in the absolute value. The gcd of the mentioned $m \times m$ minors of A is 1. This shows $\gcd(A) = 1$.

To show $\mathsf{ILC}(A) = \mathrm{mincov}(\mathcal{S}) + m$, we observe that $\mathsf{ILC}(A)$ is the smallest cardinality of $\gamma \subseteq [m + \ell]$, for which $\gcd(A_\gamma) = 1$ holds. Every such sub-matrix A_γ of A contains the first m columns of A. Indeed, if the column $-p_i\boldsymbol{e}_i$ is missing, then the i-th row of A_γ is divisible by q, which implies that $\gcd(A_\gamma)$ is divisible by q, a contradiction. Consider the sub-family $\mathcal{S}'$ of $\mathcal{S}$, consisting of all $S \in \mathcal{S}$, for which the column $q\chi_S$ occurs in A_γ. The sub-family $\mathcal{S}'$ covers $[m]$. Indeed, otherwise there would exist an $i \in [m]$ not covered by any element of $\mathcal{S}'$. But then, the i-th row of

A_γ would be divisible by p_i, which would imply that $\gcd(A_\gamma)$ is divisible by p_i, a contradiction. The above arguments show $\gcd(A) = 1$ and $\mathsf{ILC}(A) = \mathrm{mincov}(\mathcal{S}) + m$.

The equality $\gcd(A) = 1$ can be phrased as $\mathcal{L}(A) = \mathbb{Z}^m$. Since the columns of A positively span $\mathbb{R}^m$, we obtain $\mathrm{Sg}(A) = \mathcal{L}(A) = \mathbb{Z}^m$. The equality $\mathrm{Sg}(A) = \mathcal{L}(A)$ implies $\mathsf{ICC}(A) \geq \mathsf{ILC}(A)$. To see that the equality $\mathsf{ICC}(A) = \mathsf{ILC}(A)$ is true, just observe that for every matrix A_γ satisfying $\gcd(A_\gamma) = 1$, which we have analyzed above, the columns of A_γ positively span $\mathbb{R}^m$. This means $\mathrm{Sg}(A_\gamma) = \mathcal{L}(A_\gamma) = \mathbb{Z}^m$ and implies the equality $\mathsf{ICC}(A) = \mathsf{ILC}(A)$.

From $\mathrm{Sg}(A) = \mathcal{L}(A)$ and $\mathsf{ILC}(A) = \mathsf{ICC}(A)$, we easily obtain $\mathsf{ILR}(A) \leq \mathsf{ICR}(A) \leq \mathsf{ICC}(A) = \mathsf{ILC}(A)$, because $\mathsf{ILR}(A) \leq \mathsf{ILC}(A)$, $\mathsf{ICR}(A) \leq \mathsf{ICC}(A)$ and every representation of $\boldsymbol{b} \in \mathrm{Sg}(A)$ as a non-negative integer linear combination of the columns of A is also a representation of $\boldsymbol{b} \in \mathcal{L}(A)$ as an integer linear combination of the columns of A. Thus, to conclude the proof, it suffices to verify $\mathsf{ILR}(A) \geq \mathrm{mincov}(\mathcal{S}) + m$. We use $\boldsymbol{b} := (1, \ldots, 1)^\top \in \mathbb{Z}^m$. Consider a sub-matrix A_γ of A, for which the equation $A_\gamma \boldsymbol{x} = \boldsymbol{b}$ has an integer solution $\boldsymbol{x}$. Then A_γ contains each of the m columns $-p_i \boldsymbol{e}_i$: if the column $-p_i \boldsymbol{e}_i$ is missing, then the coefficients of the left-hand side of the i-th equation of the system $A_\gamma \boldsymbol{x} = \boldsymbol{b}$ are divisible by q, while the right-hand side coefficient is 1, which contradicts the solvability of the system. Let $\mathcal{S}'$ be the sub-family of $\mathcal{S}$ consisting of those S, for which the column $q\chi_S$ occurs in A_γ. The sets of the family $\mathcal{S}'$ cover $[m]$: if some element $i \in [m]$ was not in any of the sets $S \in \mathcal{S}'$, then the coefficients on the left-hand side of the i-th equation of the system $A_\gamma \boldsymbol{x} = \boldsymbol{b}$ would be divisible by p_i, while the the right-hand side coefficient is 1, which again contradicts the solvability of the system. □

Proof of Theorem 5, Parts (iii) and (iv) We derive the strong NP-hardness of all four problems by means of Lemma 10, which helps to construct a polynomial-time reduction from the set-cover problem. Consider a family $\mathcal{S}$ of subsets of $[m]$ that cover $[m]$. We want to reduce verification of $\mathrm{mincov}(\mathcal{S}) \leq k$ to the verification of any of the four inequalities in the assertion. Our reduction is the map $\mathcal{S} \mapsto A$, described in Lemma 10, for which we fix the prime numbers $p_1, \ldots, p_m, q$ to be the first $m+1$ prime numbers. As in the first part of the proof, these prime numbers can be computed in polynomial time in the size of $\mathcal{S}$, which means that the respective map $\mathcal{S} \mapsto A$ is computable in polynomial time. Furthermore, the first $m+1$ prime numbers are of order $O(m \log_e m)$, which implies that the unary encoding of A has a polynomial size in $\mathcal{S}$. In view of Lemma 10, we obtain the desired hardness assertions, as verifying $\mathrm{mincov}(\mathcal{S}) \leq k$ is reduced to verifying $\mathsf{ILC}(A) \leq m + k$, where $\mathsf{ILC}(A) = \mathsf{ILR}(A) = \mathsf{ICC}(A) = \mathsf{ICR}(A)$.

To show that $\mathsf{ILC}(A) \leq k$ is strongly NP-complete, we need to check that this problem is in NP with A represented in the binary encoding. Clearly, a set $\gamma \subseteq [n]$ with at most k elements satisfying $\mathcal{L}(A_\gamma) = \mathcal{L}(A)$ can be used as a certificate for $\mathsf{ILC}(A) \leq k$. The verification of $\mathcal{L}(A_\gamma) = \mathcal{L}(A)$ for a given A and γ can be carried in polynomial time: it suffices to check that each column of A is in $\mathcal{L}(A_\gamma)$. Each such check can be done by solving a system of linear Diophantine equations.

The NP-completeness of $\mathsf{ICC}(A) \leq k$ is proved as in the case $m = 1$ of one row. Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ be the columns of A. The certificate consists of a set $\tau \subseteq [n]$ with at most k elements and solutions of problems $A_\tau \boldsymbol{x} = \boldsymbol{a}_i$, $\boldsymbol{x} \in \mathbb{Z}^{|\tau|}_{\geq 0}$, with $i \in [n]$, that have a polynomial description size. □

6.3 Proof of Theorem 6

Presburger arithmetic is the first-order theory of the integer numbers with addition (but no multiplication) and the usual order $\leq$. A *Presburger statement* is a quantified expression of the form

$$Q_1 x_1 \in \mathbb{Z} \cdots Q_k x_k \in \mathbb{Z} \ : \ \Phi(x_1, x_2, \ldots, x_k),$$

where $Q_1, \ldots, Q_k \in \{\forall, \exists\}$ are quantifiers over integer variables $x_1, \ldots, x_k$ and $\Phi(x_1, \ldots, x_k)$ is a Boolean combination of linear inequalities with integer coefficients in the variables $x_1, \ldots, x_k$. In the 1920's Presburger showed that there is an algorithm, based on elimination of quantifiers, to verify the validity of such statements. But it is also known that deciding general Presburger statements is much harder than deciding NP-complete problems. For example, for statements with a fixed number i of quantifier alternations that start with an existential quantifier the following is known: deciding such statements is complete for the level $\Sigma_{i-1}^{\mathrm{EXP}}$ of the exponential hierarchy for $i \geq 2$ (see [22, Sect. 5]) and complete for the level Σ_{i-2}^{P} of the polynomial hierarchy when $i \geq 3$ and, additionally, the number of variables k and the number of Boolean operations used in Φ is also fixed (see [27]).

Proof of Theorem 6 Using Hermite Normal Form of A with respect to the row transformations, we can reduce the general case to the case of A having full row rank. In particular, this means $m \leq n$. It is clear that one can express the condition $\mathsf{ILR}(A) \leq k$ and $\mathsf{ICR}(A) \leq k$ as the Presburger statement

$$\forall \boldsymbol{x} \in D^n \ \exists \boldsymbol{y} \in D^k : \bigvee_{\tau \in \binom{[n]}{k}} (A\boldsymbol{x} = A_\tau \boldsymbol{y}), \tag{20}$$

with $D = \mathbb{Z}$ and $D = \mathbb{Z}_{\geq 0}$, respectively. Note that, though in our definition of a Presburger statement the quantified variables have values in $\mathbb{Z}$, it is easy to model quantified variables from $\mathbb{Z}_{\geq 0}$ via a slight reformulation: For example: $\forall x \in \mathbb{Z}_{\geq 0} : \Phi(x)$ can also be formulated as $\forall x \in \mathbb{Z} : ((x \geq 0) \Rightarrow \Phi(x))$. When n is fixed, (20) is a so-called *short* Presburger formula, which means the number of quantified variables as well as the number of Boolean operations used in the formula are fixed. For our formula, we can assume $k \leq n$ because both $\mathsf{ILR}(A)$ and $\mathsf{ICR}(A)$ are at most n. Thus, the number of quantified variables is at most $2n$. The number of disjunctions used is at most 2^n. Each system $A\boldsymbol{x} = A_\tau \boldsymbol{y}$ is a conjunction of m equalities, which means that we have used at most $m2^n \leq n2^n$ conjunctions. It is known that short Presburger statements with one quantifier alternation are solvable in polynomial time. This is explicitly stated as Theorem 1.9 in [28], where the authors of [28] refer to the work of Woods [34] and their own work [27]. We note that the proofs from [27,34] rely on the algorithmic theory of generating functions (See [5,7,8,14]). Since the short statement (20) has one quantifier alternation we obtain the polynomial-time solvability of the problems $\mathsf{ILR}(A) \leq k$ and $\mathsf{ICR}(A) \leq k$ when n is fixed. Consequently, for computing $\mathsf{ILR}(A)$ and $\mathsf{ICR}(A)$ it suffices to check the validity of n respective short Presburger statements (20) that arise by choosing n possible $k \in [n]$.

We now show that $\mathsf{ILC}(A)$ and $\mathsf{ICC}(A)$ are computable in polynomial time, too, when n is fixed. For $\mathsf{ILC}(A)$ this is easy to see. It suffices to determine all subsets $\gamma \subseteq [n]$ for which $\mathcal{L}(A) = \mathcal{L}(A_\gamma)$ holds, where the verification of $\mathcal{L}(A) = \mathcal{L}(A_\gamma)$ can be reduced to solving n systems of Diophantine equations. $\mathsf{ILC}(A)$ is the minimum cardinality among all such subsets. For $\mathsf{ICC}(A)$, one can determine all subsets $\gamma \subseteq [n]$ for which $\mathrm{Sg}(A) = \mathrm{Sg}(A_\gamma)$ holds. For checking $\mathrm{Sg}(A) = \mathrm{Sg}(A_\gamma)$ one can check whether each column of A belongs to $\mathrm{Sg}(A_\gamma)$. Each such check is reduced to solving a feasibility problem of linear integer programming in at most n variables and thus can be done in polynomial time, as n is fixed (see Section 18.4 in [32]). □

Acknowledgements An earlier shorter version of this paper appeared in the IPCO 2020 proceedings. We are grateful to the referees for suggesting several new improvements incorporated in this version. Further, we thank Christian Elsholtz for pointing to [20]. The second author is supported by the DFG (German Research Foundation) within the Project Number 413995221. The third author gratefully acknowledges partial support from NSF DMS-Grant 1818969.

7 Appendix

Even though Theorem 7 is available in the specialized literature (see, for example, [20]), it does not seem to be well known, and we have not found any elementary group theory source containing a complete self-contained proof of this result. Thus, we prove Theorem 7 relying only on the basic facts from group theory.

The following lemma shows that by "projecting out" a subset of a non-redundant set, one obtains a smaller non-redundant set.

Lemma 11 *Let S be a non-redundant subset of a finite Abelian group G. Let $T \subseteq S$ and consider the canonical homomorphism $\phi : G \to G/\langle T\rangle$. Then $\phi(S\backslash T)$ is a non-redundant subset of $G/\langle T\rangle$ of cardinality $|S| - |T|$.*

Proof For all $x', x'' \in S\backslash T$ with $x' \neq x''$, we have $\phi(x') \neq \phi(x'')$. If the latter was not the case, then we had $x' - x'' \in \langle T\rangle$, which implies $x' \in \langle T \cup \{x''\}\rangle$. Consequently, $\langle S\rangle = \langle S\backslash\{x'\}\rangle$ which contradicts the fact that S is non-redundant. It follows that ϕ is injective on $S\backslash T$. Hence $\phi(S\backslash T)$ has $|S\backslash T| = |S| - |T|$ elements. Let us verify that $\phi(S\backslash T)$ is non-redundant subset of $G/\langle T\rangle$. If $\phi(S\backslash T)$ were redundant, then we had $\phi(S\backslash T) = \phi(U)$ for some proper subset U of S. This means $\langle S\backslash T\rangle + \langle T\rangle = \langle U\rangle + \langle T\rangle$. The latter equality can be simplified to $\langle S\rangle = \langle U \cup T\rangle$. Thus, the proper subset $U \cup T$ of S generates $\langle S\rangle$, which contradicts the non-redundancy of S.

Lemma 12 *Let G be a finite Abelian group that can be decomposed as $G = F \oplus H$, where F is cyclic and $|F| \cdot h = 0$ holds for every $h \in H$. Then $G/\langle f + h\rangle \simeq H$ holds for every generator f of F and every $h \in H$. Furthermore, the map $\phi : H \to G/\langle f + h\rangle$ that sends x to the residue class of x in the group G with respect to the equivalence modulo $\langle f + h\rangle$ is a group isomorphism.*

Proof We claim that the group $\langle f + h\rangle$ is cyclic of order $|F|$. Note that $z(f + h)$ with $z \in \{0, \ldots, |F| - 1\}$ are $|F|$ distinct elements of F, because zf with $z \in \{0, \ldots, |F| - 1\}$ are the $|F|$ distinct elements of F. On the other hand $|F| \cdot (f + h) = |F| \cdot f + |F| \cdot h = 0 + 0 = 0$, which shows that $\langle f + h\rangle$ has no other elements.

Since the order of $\langle f + h\rangle$ is $|F|$, the order of $G/\langle f + h\rangle$ coincides with the order of H. Thus, to verify the assertion, it suffices to show that ϕ is injective. To this end, we check that the kernel of ϕ is $\{0\}$. Consider an arbitrary $x \in H$ with $\phi(x) = 0$. This means, $x \in \langle f + h\rangle$. Thus, $x = z(f + h)$ holds for some $z \in \mathbb{Z}$. In view of $G = F \oplus H$, $x, h \in H$ and $f \in F$, the latter implies $0 = zf$ and $x = zh$. Since f is a generator of the cyclic group f, the equality $0 = zf$ implies that z is a multiple of $|F|$. But then $zh = 0$, which implies $x = 0$. This shows that ϕ has trivial kernel and concludes the proof. □

Proof of Theorem 7 Consider the decomposition of G into direct sum of primary cyclic groups:

$$G = \bigoplus_{i=1}^{m} \bigoplus_{j=1}^{k_i} G_{i,j}, \tag{21}$$

where $k_1, \ldots, k_m \in \mathbb{N}$ and, for each $i \in [m]$ and $j \in [k_i]$, the direct summand $G_{i,j}$ is a primary cyclic group of order $p_i^{n_{i,j}}$ with $n_{i,j} \in \mathbb{N}$. One has $\kappa(G) = \sum_{i=1}^{m} k_i$. See Chapter 5 in [15] for details. Some algebra books call this the elementary divisor decomposition of G.

First note that G contains a non-redundant set of cardinality $\kappa(G)$, which can be constructed by picking a generator of $G_{i,j}$ for each cyclic group $G_{i,j}$ from the decomposition of G. We thus need to show that for an arbitrary non-redundant subset S of G the inequality $|S| \leq \kappa(G)$ is fulfilled. We argue by induction on $\Omega(|G|)$. If $\Omega(|G|) = 1$, then $|G|$ is prime. Consequently, G is a cyclic group of prime order, which means that every non-zero element of G generates the whole G. We conclude that $|S| \leq 1 = \kappa(G)$.

We fix an arbitrary integer $N > 1$ and assume that the bound on the cardinality of non-redundant sets is true in every finite Abelian group $\tilde{G}$ with $\Omega(|\tilde{G}|) < N$. Let G be a finite Abelian group with $\Omega(|G|) = N$. We verify the bound for the group G.

Using projection homomorphisms $x \mapsto x_{i,j}$ from G to $G_{i,j}$, each $x \in G$ can be uniquely written as

$$x = \sum_{i=1}^{m} \sum_{j=1}^{k_i} x_{i,j}.$$

Case 1: There exist $i \in [m]$ and $j \in [k_i]$ such that for every $x \in S$, the group $\langle x_{i,j}\rangle$ is a proper subgroup of $G_{i,j}$. Assume, without loss of generality, that $\langle x_{1,1}\rangle$ is a proper

subgroup of $G_{1,1}$ for every $x \in S$. Since $G_{1,1}$ is cyclic of order $p_1^{n_{1,1}}$, every subgroup properly contained in $G_{1,1}$ is a subgroup of the (unique) cyclic subgroup $\tilde{G}_{1,1}$ of $G_{1,1}$ of order $p_1^{n_{1,1}-1}$. Fix $\tilde{G}_{i,j} := G_{i,j}$ for all $i \in [m]$ and $j \in [k_i]$ with $(i, j) \neq (1, 1)$. It follows that S is a subset of the group

$$\tilde{G} = \bigoplus_{i=1}^{m} \bigoplus_{j=1}^{k_i} \tilde{G}_{i,j},$$

where $\Omega(|\tilde{G}|) = \Omega(|G|) - 1 < N$. Thus, by the induction assumption we obtain $|S| \leq \kappa(\tilde{G})$. Since $\kappa(\tilde{G}) \leq \kappa(G)$ we conclude $|S| \leq \kappa(|G|)$.

Case 2: For all $i \in [m],\ j \in [k_i]$, there exists some $x \in S$ such that $\langle x_{i,j} \rangle = G_{i,j}$. Without loss of generality, we can assume that the numbers $n_{i,j}$ are ordered so that

$$n_{i,1} \geq \cdots \geq n_{i,k_i} \tag{22}$$

holds for every $i \in [m]$. We represent G as the direct sum $G = F \oplus H$, where

$$F := \bigoplus_{i=1}^{m} G_{i,1} \qquad \text{and} \qquad H := \bigoplus_{i=1}^{m} \bigoplus_{j=2}^{k_i} G_{i,j}.$$

Since the orders of the cyclic groups $G_{1,1}, \ldots, G_{m,1}$ are pairwise relatively prime, the Chinese Remainder Theorem implies that F is a cyclic group of order $d := \prod_{i=1}^{m} p_{i,1}^{n_{i,1}}$. Each $x \in G$ can be projected onto F and H. That is, for $x \in G$, we introduce

$$x_F := \sum_{i=1}^{m} x_{i,1}, \qquad \text{and} \qquad x_H := \sum_{i=1}^{m} \sum_{j=2}^{k_i} x_{i,j}.$$

Choose a subset T of S of cardinality at most m by picking, for each $i \in \{1, \ldots, m\}$, an element x of S satisfying $\langle x_{i,1} \rangle = G_{i,1}$. Since F is cyclic, the order of $\langle \{x_F : x \in T\} \rangle$ is the least common multiplier of the orders of $G_{1,1}, \ldots, G_{m,1}$ and thus is equal to $|F|$. This means $\langle \{x_F : x \in T\} \rangle = F$.

We fix $t \in \langle T \rangle$ such that t_F is a generator of the cyclic group F and consider the projection homomorphism

$$\psi : G \to G/\langle t \rangle .$$

In view of (22), the decomposition $G = F \oplus H$ satisfies the assumptions of Lemma 12. This implies that $\phi := \psi|_H$ is an isomorphism from H to $G/\langle t \rangle$.

Lemma 11 implies that $S' = \psi(S \backslash T)$ is a non-redundant subset of $G/\langle t \rangle$ of cardinality $|S| - |T|$. Consequently, using the induction assumption for $G/\langle t \rangle \simeq H$, we

obtain the desired bound

$$|S| = \underbrace{|S'|}_{\leq \kappa(G/\langle t\rangle)} + \underbrace{|T|}_{\leq m} \leq \kappa(G/\langle t\rangle) + m = \kappa(H) + m = \kappa(G)$$

on the cardinality of $|S|$. □

References

1. Aliev, I., Averkov, G., De Loera, J.A., Oertel, T.: Optimizing sparsity over lattices and semigroups. In: Integer Programming and Combinatorial Optimization, Volume 12125 of Lecture Notes in Comput. Sci., pp. 40–51. Springer, Cham (2020)
2. Aliev, I., De Loera, J.A., Eisenbrand, F., Oertel, T., Weismantel, R.: The support of integer optimal solutions. SIAM J. Optim. **28**(3), 2152–2157 (2018)
3. Aliev, I., De Loera, J.A., Oertel, T., O'Neill, C.: Sparse solutions of linear diophantine equations. SIAM J. Appl. Algebra Geom. **1**(1), 239–253 (2017)
4. Averkov, G., Chavez, A., De Loera, J.A., Gillespie, B.: The lattice of cycles of an undirected graph. Linear Algebra Appl. **611**, 213–236 (2021)
5. Baldoni, V., Berline, N., De Loera, J., Dutra, B., Köppe, M., Moreinis, S., Pinto, G., Vergne, M., Wu, J.: A User's Guide for LattE Integrale v1.7.2 (2014). http://www.math.ucdavis.edu/~latte/
6. Barvinok, A.I.: A Course in Convexity. Graduate Studies in Mathematics, vol. 54. American Mathematical Society, Providence (2002)
7. Barvinok, A.I., Pommersheim, J.E.: An algorithmic theory of lattice points in polyhedra. In: New Perspectives in Algebraic Combinatorics (Berkeley, CA, 1996–97), Volume 38 of Math. Sci. Res. Inst. Publ., pp. 91–147. Cambridge Univ. Press, Cambridge (1999)
8. Barvinok, A.I., Woods, K.: Short rational generating functions for lattice point problems. J. AMS **16**(4), 957–979 (2003)
9. Boche, H., Calderbank, R., Kutyniok, G., Vybíral, J.: A survey of compressed sensing. In: Compressed Sensing and Its Applications, Appl. Numer. Harmon. Anal., pp. 1–39. Birkhäuser/Springer, Cham (2015)
10. Candès, E., Rudelson, M., Tao, T., Vershynin, R.: Error correction via linear programming. In: 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05), pp. 668–681 (2005)
11. Candès, E.J., Romberg, J.K., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math. **59**(8), 1207–1223 (2006)
12. Candès, E.J., Tao, T.: Decoding by linear programming. IEEE Trans. Inform. Theory **51**(12), 4203–4215 (2005)
13. Cioabă, S.M., Cameron, P.J.: A graph partition problem. Am. Math. Mon. **122**(10), 972–982 (2015)
14. De Loera, J.A., Hemmecke, R., Tauzer, J., Yoshida, R.: Effective lattice point counting in rational convex polytopes. J. Symb. Comput. **38**(4), 1273–1302 (2004)
15. Dummit, D.S., Foote, R.M.: Abstract Algebra, 3rd edn. Wiley, New York (2004)
16. Eisenbrand, F., Shmonin, G.: Carathéodory bounds for integer cones. Oper. Res. Lett. **34**(5), 564–568 (2006)
17. Flinth, A., Kutyniok, G.: PROMP: a sparse recovery approach to lattice-valued signals. Appl. Comput. Harmon. Anal. **45**(3), 668–708 (2018)
18. Fukshansky, L., Needell, D., Sudakov, B.: An algebraic perspective on integer sparse recovery. Appl. Math. Comput. **340**, 31–42 (2019)
19. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman and Co., New York (1979)
20. Geroldinger, A., Halter-Koch, F.: Non-Unique Factorizations: Algebraic, Combinatorial and Analytic Theory. Pure and Applied Mathematics. Chapman and Hall/CRC, Boca Raton (2006)
21. Gruber, P.M., Lekkerkerker, C.G.: Geometry of Numbers. North-Holland Mathematical Library, vol. 37, 2nd edn. North-Holland Publishing Co., Amsterdam (1987)
22. Haase, C.: A survival guide to Presburger arithmetic. ACM SIGLOG News **5**(3), 67–82 (2018)

23. Hardy, G.H., Wright, E.M., Heath-Brown, R., Silverman, J.: An Introduction to the Theory of Numbers. Oxford Mathematics. OUP, Oxford (2008)
24. Konyagin, S.V.: On the recovery of an integer vector from linear measurements. Mat. Zametki **104**(6), 863–871 (2018)
25. Lovász, L.: Matching structure and the matching lattice. J. Comb. Theory Ser. B **43**(2), 187–222 (1987)
26. Natarajan, B.K.: Sparse approximate solutions to linear systems. SIAM J. Comput. **24**(2), 227–234 (1995)
27. Nguyen, D., Pak, I.: Complexity of short Presburger arithmetic. In: Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, pp. 812–820 (2017)
28. Nguyen, D., Pak, I.: Short Presburger arithmetic is hard. SIAM J. Computi., (0):STOC17–1 (2019)
29. Oertel, T., Paat, J., Weismantel, R.: Sparsity of integer solutions in the average case. In: Integer Programming and Combinatorial Optimization, Volume 11480 of Lecture Notes in Comput. Sci., pp. 341–353. Springer, Cham (2019)
30. Oertel, T., Paat, J., Weismantel, R.: The distributions of functions related to parametric integer optimization. SIAM J. Appl. Algebra Geom. **4**(3), 422–440 (2020)
31. Rossi, M., Haimovich, A.M., Eldar, Y.C.: Spatial compressive sensing for MIMO radar. IEEE Trans. Signal Process. **62**(2), 419–430 (2014)
32. Schrijver, A.: Theory of Linear and Integer Programming. Wiley-Interscience Series in Discrete Mathematics. Wiley, Chichester (1986). A Wiley-Interscience Publication
33. Vaaler, J.D.: A geometric inequality with applications to linear forms. Pac. J. Math. **83**(2), 543–553 (1979)
34. Woods, K.: Presburger arithmetic, rational generating functions, and quasi-polynomials. J. Symb. Log. **80**(2), 433–449 (2015)